

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 10, 2004, 08:26:09 ; Search time 4183 Seconds
(without alignments)
10258.093 Million cell updates/sec

Title: US-09-972-032-1

Perfect score: 990

Sequence: 1 ggaattgtctcagagccaa.....aaatgagcgccgaagt 990

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:

```

1: gb_ba:*
2: gb_btg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:

```

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	797.6	80.6	182230	9	AC135050	AC135050 Homo sapi
2	792.8	80.1	195476	2	AC135044	AC135044 Homo sapi
C 3	87.4	8.8	1602	6	AX884081	AX884081 Sequence
C 4	87.4	8.8	1602	6	BD160707	BD160707 Primer fo
C 5	87.4	8.8	1602	9	AK024161	AK024161 Homo sapi
C 6	87.4	8.8	32360	9	AC005339	AC005339 Homo sapi
7	86.6	8.7	176735	2	AC021271	AC021271 Homo sapi
8	86	8.7	185316	2	AC134386	AC134386 Papio anu
C 9	85.8	8.7	2209	9	BC017672	BC017672 Homo sapi
10	85.8	8.7	80429	2	AC010603	AC010603 Homo sapi
C 11	85.8	8.7	161267	9	AC008378	AC008378 Homo sapi
12	85.8	8.7	165762	2	AC124263	AC124263 Homo sapi
13	85.8	8.7	185378	9	AC118754	AC118754 Homo sapi
14	85.8	8.7	190466	2	AC022285	AC022285 Homo sapi
15	85.8	8.7	199463	2	AC129497	AC129497 Homo sapi
C 16	85.6	8.6	101658	9	AL590005	AL590005 Human DNA
17	85.6	8.6	174578	9	AC013287	AC013287 Homo sapi
18	85.4	8.6	143065	9	HS20208	AL031848 Human DNA
19	85.4	8.6	173028	9	AC090371	AC090371 Homo sapi
C 20	85.4	8.6	193486	9	AC012123	AC012123 Homo sapi
C 21	85.2	8.6	119841	9	AC005002	AC005002 Homo sapi
22	85.2	8.6	123291	9	AC104070	AC104070 Homo sapi
23	85.2	8.6	174133	9	AC093183	AC093183 Homo sapi
C 24	85.2	8.6	176875	9	AC018638	AC018638 Homo sapi
C 25	85	8.6	60500	9	AC131238	AC131238 Homo sapi
C 26	85	8.6	91692	9	AC084253	AC084253 Homo sapi
C 27	85	8.6	94296	2	AL160261	AL160261 Homo sapi
C 28	85	8.6	123556	9	AL391863	AL391863 Human DNA
C 29	85	8.6	159933	2	AC009692	AC009692 Homo sapi
C 30	85	8.6	161970	9	AP005660	AP005660 Homo sapi
31	85	8.6	187268	9	AC094022	AC094022 Homo sapi
32	85	8.6	200598	2	AC012505	AC012505 Homo sapi
C 33	85	8.6	205523	9	CNS01RH1	AL160237 Human chr
C 34	85	8.6	214866	9	AC069544	AC069544 Homo sapi
C 35	84.8	8.6	114596	9	HS1063H2	AL035683 Human DNA
C 36	84.8	8.6	169197	9	AC129096	AC129096 Papio anu
37	84.8	8.6	202399	9	AC090966	AC090966 Papio anu
C 38	84.8	8.6	231260	2	AL160172	AL160172 Homo sapi
C 39	84.6	8.5	91767	9	AC007748	AC007748 Homo sapi
C 40	84.6	8.5	114950	9	AC097652	AC097652 Homo sapi
C 41	84.6	8.5	134308	9	AC011488	AC011488 Homo sapi
42	84.6	8.5	159440	2	AC079012	AC079012 Homo sapi
43	84.6	8.5	171611	2	AC119423	AC119423 Papio anu
44	84.6	8.5	183301	2	AC092487	AC092487 Homo sapi
45	84.6	8.5	210608	9	AC006028	AC006028 Homo sapi

ALIGNMENTS

```

RESULT 1
AC135050
LOCUS      AC135050      182230 bp      DNA      linear      PRI 27-FEB-2003
DEFINITION Homo sapiens chromosome 16 clone RP11-196G11, complete sequence.
ACCESSION  AC135050
VERSION    AC135050.3   GI:28570306
KEYWORDS   HTG.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 182230)
AUTHORS   DOE Joint Genome Institute, Stanford Human Genome Center and Los
            Alamos National Laboratory.
TITLE     Direct Submission

```

```

JOURNAL Unpublished
REFERENCE 2 (bases 1 to 182230)
AUTHORS DOE Joint Genome Institute.
TITLE Direct Submission
JOURNAL Submitted (05-OCT-2002) Production Sequencing Facility, DOE Joint
        Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE 3 (bases 1 to 182230)
AUTHORS DOE Joint Genome Institute.
TITLE Direct Submission
JOURNAL Submitted (22-OCT-2002) Production Sequencing Facility, DOE Joint
        Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE 4 (bases 1 to 182230)
AUTHORS DOE Joint Genome Institute, Stanford Human Genome Center and Los
        Alamos National Laboratory.
TITLE Direct Submission
JOURNAL Submitted (27-FEB-2003) DOE Joint Genome Institute, 2800 Mitchell
        Drive, Walnut Creek, CA 94598, USA
COMMENT On Feb 27, 2003 this sequence version replaced gi:24211094.
        Draft Sequence Produced by DOE Joint Genome Institute
        www.jgi.doe.gov
        Finishing Completed at Stanford Human Genome Center and Los Alamos
        National Laboratory
        www.sngc.stanford.edu
        Quality: Phrap Quality >=40 99.7% of Sequence;
        Estimated Total Number of Errors is 0.3.
FEATURES             source
     location/Qualifiers
     1..182230
     /organism="Homo sapiens"
     /mol_type="genomic DNA"
     /db_xref="taxon:9606"
     /chromosome="16"
     /clone="RP11-196G11"

ORIGIN
Query Match      80.6%; Score 797.6; DB 9; Length 182230;
Best Local Similarity 99.3%; Pred. No. 1.4e-181;
Matches 822; Conservative 0; Mismatches 4; Indels 2; Gaps 2;

QY 126 TTTAAAGGGCCCTGCTGTTCTCCGCTCCGCTCCAGTCCCAAGTCTGGCGGG 185
Db 98012 TCTTACAGGCCCTGCTGTTCTCCGCTCCGCTCCAGTCCCAAGTCTGGCGGG 98071
QY 186 CTCAGTCCCGGTTGCCATGTGCGAGACCGCTCCGCTAAGCGTGGATGTGGCTTCG 245
Db 98072 CTCAGTCCCGGTTGCCATGTGCGAGACCGCTCCGCTAAGCGTGGATGTGGCTTCG 98131
QY 246 CTGATGCACATTGACCGGGCTCTGGACTGGCTAGGGGAAGGCGAGGCGGGAATTG 305
Db 98132 CTGATGCACATTGACCGGGCTCTGGACTGGCTAGGGGAAGGCGAGGCGGGAATTG 98191
QY 306 GSCCGGAGGCGAGGCTCCGCGACCCCGACTCGCTCCGCTCCGCTCCGCGAGCGCCT 365
Db 98192 GSCCGGAGGCGAGGCTCCGCGACCCCGACTCGCTCCGCTCCGCTCCGCGAGCGCCT 98251
QY 366 CCCGGTGGCCCTGGAGTGCAGGTCTTACCGTCCGAGATCGTCCGCAACTGGGCGAGTGT 425
Db 98252 CCCGGTGGCCCTGGAGTGCAGGTCTTACCGTCCGAGATCGTCCGCAACTGGGCGAGTGT 98311
QY 426 GCATGGGCGTGGCTAAGGCGGTGTTGTTACGATTGCCAGCGGACCTTAAGTGTTCG 485
Db 98312 GCATGGGCGTGGCTAAGGCGGTGTTGTTACGATTGCCAGCGGACCTTAAGTGTTCG 98371
QY 486 TCTCTGAAGACATGACATTAGTCTGAGGGTCTTGAAGAGTGAATCCCGCCGCCACCA 545
Db 98372 TCTCTGAAGACATGACATTAGTCTGAGGGTCTTGAAGAGTGAATCCCGCCGCCACCA 98431
QY 546 TCAATGGCGCTTAGTCTAGGAAGCGGTGTGGTGGGCGCTTAGGCGAGCGGAGAC 605
Db 98432 TCAATGGCGCTTAGTCTAGGAAGCGGTGTGGTGGGCGCTTAGGCGAGCGGAGAC 98491
QY 606 ATACCCCGAAGTGGTGGATTGTATACCGCAAGGGGTGGATCGAACCCCGCAAGACAC 665
Db 98492 ATACCCCGAAGTGGTGGATTGTATACCGCAAGGGGTGGATCGAACCCCGCAAGACAC 98551

```

666 TGGAGGCTGTGCTGCTGAGGAGGCGCCGGCA-ATCCAGTGTGCTGCTGGCTTTACAGGA 724
 98552 TGGAGGCTGTGCTGCTGAGGAGGCGCCGGCAATCCAGTGTGCTGCTGGCTTTACAGGA 98611
 725 AAGAGCTCCACCTTCT-TGGAGTGTGCAGATCGCATCTAGGTGTGTCCACCCGATGGGAG 783
 98612 AAGAGCTCCACCTTCTCTGGAGTGTGCAGATCGCATCTAGGTGTGTCCACCCGATGGGAG 98671
 784 CTGCGGCGCGGCACATGCTGCCCGCAGTACAAAGCTGATTTGGACCTGGGCGCTCTGGAC 843
 98672 CTGCGGCGCGGCACATGCTGCCCGCAGTACAAAGCTGATTTGGACCTGGGCGCTCTGGAC 98731
 844 TTCCCTGATTTCTGCTTGTGCAATCTCAGCAAAAGTCTCTCCGTTGGTGGCTCTTCATCCA 903
 98732 TTCCCTGATTTCTGCTTGTGCAATCTCAGCAAAAGTCTCTCCGTTGGTGGCTCTTCATCCA 98791
 904 CTCTCTCACTTCTGCTTGTGCAATCTCAGCAAAAGTCTCTCCGTTGGTGGCTCTTCATCCA 951
 98792 CTCTCTCACTTCTGCTTGTGCAATCTCAGCAAAAGTCTCTCCGTTGGTGGCTCTTCATCCA 98839

RESULT 2
 AC135044 195476 bp DNA linear HTG 05-OCT-2002
 LOCUS Homo sapiens chromosome 16 clone CTD-2551B20, WORKING DRAFT
 DEFINITION SEQUENCE, 22 unordered pieces.
 AC135044
 AC135044.1 GI:23505535
 VERSION HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ACTIVEFIN.
 KEYWORDS Homo sapiens (human)
 SOURCE Homo sapiens
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 195476)
 AUTHORS DOE Joint Genome Institute.
 TITLE Sequencing of Human Chromosome 16
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 195476)
 AUTHORS DOE Joint Genome Institute.
 TITLE Direct Submission
 JOURNAL Submitted (05-OCT-2002) Production Sequencing Facility, DOE Joint
 Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
 COMMENT -----Genome Center
 Center: Joint Genome Institute
 Center Code: JGI
 Web site: http://www.jgi.doe.gov

 Project Information
 Center Project Name: 809609
 Center clone name: CITB-El_2551B20

 Summary Statistics
 Consensus quality: 171229 bases at least Q40
 Consensus quality: 182638 bases at least Q30
 Consensus quality: 188095 bases at least Q20
 Estimated insert size: 170000; agarose-fp estimation
 Estimated insert size: 193376; sum-of-contigs estimation
 Quality coverage: 7.63 in Q20 bases; agarose-fp estimation
 Quality coverage: 6.7 in Q20 bases; sum-of-contigs estimation.
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 22 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.
 * 1 1202: contig of 1202 bp in length
 * 1203 1302: gap of unknown length
 * 1303 2468: contig of 1166 bp in length
 * 2469 2568: gap of unknown length
 * 2569 4077: contig of 1509 bp in length


```

100% identity.-(25061..24842) A477268 zu43c12.r1 Soares
ovary tumor N80T Homo sapiens cDNA clone 740758 5';
(350..570); 98% identity.-(24792..25272) N32339 yw82g09.s1
Homo sapiens cDNA clone 258782 3'; Score: 925 Identity:
476/480 (99%)..(25329..25028) AA579149 nf28a04.s1
NCI_CGAP_Pri Homo sapiens cDNA clone IMAGE:915054;
(134..434); 99% identity.-(24790..25061) AA477269
zu43c12.s1 Soares ovary tumor N80T Homo sapiens cDNA
clone 740758 3'; (272..1); 100% identity.-Additional EST
matches:
AA581955, AA467935, AI038745, AI041764, T24716"
Complement[Joan(25250..25329,27176..27248,27914..27995,
32121..32185)]
/Note="Hypothetical partial human protein"
/codon_start=1
/evidence=not experimental
/product="R33729.1"
/protein_id="AAC37824.1"
/db_xref="GI:3355455"

Query Match      8.8%; Score 87.4; DB 9; Length 32360;
Best Local Similarity 85.8%; Pred. NO. 2e-10;
Matches 97; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 34 CGATCTCTGACCTGATCGCGCCGCTCAGCTTCCCAAGGCTGGGATTACAGCA 93
Db 20347 CGATCTCTGACCTGATCGCGCCGCTCAGCTTCCCAAGGCTGGGATTACAGCG 20288

QY 94 TGAGCCACCGCGCGCCGCGCCGACACCTAGTTTAAAGGCGCCCTGCTGTG 146
Db 20287 TGAGCCACCGCGCCGCTGCGCCGAGCGCCCTTTTAAAGGACCCAGTCCTTG 20235

RESULT 7
AC021271
LOCUS
DEFINITION Homo sapiens clone RP11-15H15, WORKING DRAFT SEQUENCE, 31 unordered
pieces.
AC021271
AC021271.5 GI:10047755
VERSION
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 176735)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Aldrich,E., Allen,N.,
Anderson,S., Baldwin,J., Barna,N., Beckerly,R., Bida,F.,
Boguslavskiy,L., Boukhgalter,B., Brown,A., Burkett,G., Castle,A.,
Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P.,
Dearellano,K., Dewar,K., Domino,M., Doyle,M., Fenebor,J.,
Ferreira,P., FitzHugh,W., Forrest,C., Gage,D., Galagan,J.,
Gardyna,S., Grant,G., Hages,B., Heaford,A., Horton,L.,
Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
Landers,T., Lehoczy,J., Levine,R., Lie,C., Liu,G., Locke,K.,
Macdonald,P., Margulis,N., McEwan,P., McGurk,A., McKernan,K.,
McPheters,R., Meldrim,J., Meneus,L., Morrow,J., Naylor,J.,
Nowman,C.H., O'Connor,T., O'Donnell,P., Oliver,T.M., Peterson,K.,
Pierre,N., Pisan,C., Pollara,V., Raymond,C., Riley,R., Rothman,D.,
Roy,A., Santos,R., Severy,P., Spencer,B., Stange-Thomann,N.,
Stojanovic,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Tirrell,A., Vassiliev,H., Viel,R., Vo,A., Wu,X., Wyman,D., Ye,W.J.,
Zimmer,A. and Zody,M.
Direct Submission
Submitted (16-JAN-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Sep 10, 2000 this sequence version replaced gi:7387323.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)

```

```

http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIGR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submission@genome.wi.mit.edu
----- Project Information
Center project name: L3461
Center clone name: 15_H_15
----- Summary Statistics
Sequencing vector: M13; W7815; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 149852 bases at least Q40
Consensus quality: 164690 bases at least Q30
Consensus quality: 170387 bases at least Q20
Insert size: 176000; agarose-fp
Insert size: 173735; sum-of-contigs
Quality coverage: 3.3 in Q20 bases; agarose-fp
Quality coverage: 3.4 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 31 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 1130: contig of 1130 bp in length
* 1131 1230: gap of 100 bp
* 1231 2544: contig of 1314 bp in length
* 2545 2544: gap of 100 bp
* 2545 3750: contig of 1106 bp in length
* 3751 3850: gap of 100 bp
* 3851 4962: contig of 1111 bp in length
* 4962 5061: gap of 100 bp
* 5062 6442: contig of 1381 bp in length
* 6443 6542: gap of 100 bp
* 6543 7884: contig of 1342 bp in length
* 7885 7884: gap of 100 bp
* 7885 9320: contig of 1336 bp in length
* 9321 9420: gap of 100 bp
* 9421 10821: contig of 1401 bp in length
* 10822 10921: gap of 100 bp
* 10922 12048: contig of 1127 bp in length
* 12049 12148: gap of 100 bp
* 12149 13179: contig of 1031 bp in length
* 13180 13279: gap of 100 bp
* 13280 15318: contig of 2039 bp in length
* 15319 15418: gap of 100 bp
* 15419 19113: contig of 3695 bp in length
* 19114 19213: gap of 100 bp
* 19214 22184: contig of 2971 bp in length
* 22185 22285: gap of 100 bp
* 22285 26041: contig of 3757 bp in length
* 26042 26141: gap of 100 bp
* 26142 29847: contig of 3706 bp in length
* 29848 29947: gap of 100 bp
* 29948 34786: contig of 4839 bp in length
* 34787 34886: gap of 100 bp
* 34887 38790: contig of 3904 bp in length
* 38791 38890: gap of 100 bp
* 38891 43461: contig of 4571 bp in length
* 43462 43561: gap of 100 bp
* 43562 47459: contig of 3898 bp in length
* 47460 47559: gap of 100 bp
* 47560 53268: contig of 5729 bp in length
* 53269 53388: gap of 100 bp
* 53389 59416: contig of 5928 bp in length
* 59417 64434: contig of 5018 bp in length
* 64435 64534: gap of 100 bp

```

```

* 64535 70699: contig of 6165 bp in length
* 70700 70799: gap of 100 bp
* 70800 70800: contig of 7201 bp in length
* 78001 78100: gap of 100 bp
* 78101 85100: contig of 7000 bp in length
* 85101 85200: gap of 100 bp
* 85201 94940: contig of 9740 bp in length
* 94941 95040: gap of 100 bp
* 95041 105428: contig of 10388 bp in length
* 105429 105529: gap of 100 bp
* 105529 117875: contig of 12346 bp in length
* 117875 117974: gap of 100 bp
* 117975 132905: contig of 14931 bp in length
* 132906 133005: gap of 100 bp
* 133006 150288: contig of 17283 bp in length
* 150289 150389: gap of 100 bp
* 150389 176735: contig of 26347 bp in length.
FEATURES
  source
    Location/Qualifiers
      1..176735
        /organism="Homo sapiens"
        /mol_type="genomic DNA"
        /db_xref="taxon:9606"
        /clone_lib="RPC1-11 Human Male BAC"
      1..11130
        /note="assembly_fragment
        clone end:SP6
        vector side:left"
      1231..2544
        /note="assembly_fragment"
      2645..3750
        /note="assembly_fragment"
      3851..4961
        /note="assembly_fragment"
      5062..6442
        /note="assembly_fragment"
      6543..7884
        /note="assembly_fragment"
      7985..9320
        /note="assembly_fragment"
      9421..10821
        /note="assembly_fragment"
      10922..12048
        /note="assembly_fragment"
      12149..13179
        /note="assembly_fragment"
      13280..15318
        /note="assembly_fragment"
      15419..19113
        /note="assembly_fragment"
      19214..22184
        /note="assembly_fragment"
      22285..26041
        /note="assembly_fragment"
      26142..29847
        /note="assembly_fragment"
      29948..34786
        /note="assembly_fragment"
      34887..38790
        /note="assembly_fragment"
      38891..43461
        /note="assembly_fragment"
      43562..47459
        /note="assembly_fragment"
      47560..53288
        /note="assembly_fragment"
      53389..59316
        /note="assembly_fragment"
      59417..64434
        /note="assembly_fragment"
      64535..70699
        /note="assembly_fragment"
      70800..78000

```

```

      /note="assembly_fragment"
      78101..85100
      /note="assembly_fragment"
      85201..94940
      /note="assembly_fragment"
      95041..105428
      /note="assembly_fragment"
      105529..117874
      /note="assembly_fragment"
      117975..132905
      /note="assembly_fragment"
      133006..150288
      /note="assembly_fragment"
      150389..176735
      /note="assembly_fragment"
    ORIGIN
    Query Match      8.7%; Score 86.6; DB 2; Length 176735;
    Best Local Similarity 91.1%; Pred. No. 3.2e-10;
    Matches 92; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
    QY 34 CGATCTCTGACCTCGTGCCTCGCGGCTCAGCTTCCCAAAGTCTGGGATTACAGCA 93
    Db 42019 CGATCTCTGACCTCGTGCCTCGCGGCTCAGCTTCCCAAAGTCTGGGATTACAGCT 42078
    QY 94 TGAGCCACCGCGCGCGCGCCGACACCTAGTTTAAAGGG 134
    Db 42079 TGAGCCACCGCGCGCGCGCCGACCTGACAGCAGTTTAAATGG 42119
    RESULT 8
    AC134386 185316 bp DNA linear HTG 11-SEP-2003
    LOCUS
    DEFINITION
      Papio anubis clone rp41-10k6, WORKING DRAFT SEQUENCE, 15 ordered
      pieces.
    ACCESSION
      AC134386
    VERSION
      AC134386.5 GI:31193964
    KEYWORDS
      HTG; HTGS_PHASE2; HTGS_DRAFT.
    SOURCE
      Papio anubis (olive baboon)
    ORGANISM
      Papio anubis
      Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
      Cercopithecinae; Papio.
    REFERENCE
      1 (bases 1 to 185316)
      Lau,C.C.Y., Shaikh,T. and Roe,B.A.
      Papio anubis BAC Clone rp41-10k6
      JOURNAL
      Unpublished
      2 (bases 1 to 185316)
      Lau,C.C.Y. and Roe,B.A.
      Direct Submission
      TITLE
      Submitted (26-SEP-2002) Department Of Chemistry And Biochemistry,
      The University Of Oklahoma, 620 Farrington Oval, Room 208, Norman,
      OK 73019, USA
      3 (bases 1 to 185316)
      Lau,C.C.Y., Shaikh,T. and Roe,B.A.
      Direct Submission
      TITLE
      Submitted (11-SEP-2003) Department Of Chemistry And Biochemistry,
      The University Of Oklahoma, 620 Farrington Oval, Room 208, Norman,
      OK 73019, USA
    COMMENT
      ----- Genome Center
      Center: Department Of Chemistry And Biochemistry
      The University Of Oklahoma
      Center code:UONOR
      -----
      * NOTE: This is a 'working draft' sequence. It currently
      * consists of 15 contigs. Gaps between the contigs
      * are represented as runs of N. The order of the pieces
      * is believed to be correct as given, however the sizes
      * of the gaps between them are based on estimates that have
      * been provided by the submittor.
      * This sequence will be replaced
      * by the finished sequence as soon as it is available and

```

* the accession number will be preserved.

```

* 1 2467: contig of 2467 bp in length
* 2468 2567: gap of unknown length
* 2568 5707: contig of 3140 bp in length
* 5708 5807: gap of unknown length
* 5808 8278: contig of 2471 bp in length
* 8279 8378: gap of unknown length
* 8379 11812: contig of 3434 bp in length
* 11813 17999: gap of unknown length
* 17999 25937: contig of 7938 bp in length
* 25938 37583: gap of unknown length
* 37584 51903: contig of 11546 bp in length
* 51904 52003: gap of unknown length
* 52003 62300: contig of 10297 bp in length
* 62301 74061: contig of 11661 bp in length
* 74062 85046: gap of unknown length
* 85047 92131: contig of 10885 bp in length
* 92132 120762: gap of unknown length
* 120763 150205: contig of 21531 bp in length
* 150206 150305: gap of unknown length
* 150306 185316: contig of 35011 bp in length.

```

FEATURES

```

source
1..185316
  /organism="Papio anubis"
  /mol_type="genomic DNA"
  /db_xref="taxon:9555"
  /clone_lib="RP41-10k6"
  /clone_lib="RPC1 - 41 Male (Olive) Baboon BAC Library"

```

ORIGIN

```

Query Match      8.7%; Score 86; DB 2; Length 185316;
Best Local Similarity 68.4%; Pred. No. 4.5e-10;
Matches 119; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

QY 34 CGATCTCTGACCTGCTGATCGCGCGCCCTCAGCTCCCAAGTCTGGGATTACAGGCA 93
DB 174411 CGATCTCTGACCTGCTGATCGCGCGCCCTCAGCTCCCAAGTCTGGGATTACAGGCG 174470

QY 94 TGAGCCACCGCGCGCGCCCGACACACTAGTTTAAAGGGCCCTGCTGCTGCGCGC 153
DB 174471 TAAGCCACCGCGCGCGCCCTCCTAATGCTATTTTAAAGCTCATCAGCTATCTAGCGT 174530

QY 154 TGGCGCGCTCCAGCTGCGCGAGTCTGGCGGCTCAGTCCCGGCTGCGGATGCGCATGTG 207
DB 174531 TAGTGTATTATGCTGCGCGCCAGACAAATTTCTTCTTCAATGCGCGCCAGGG 174584

```

RESULT 9

```

BC017672/c
LOCUS BC017672 2209 bp mRNA linear PRI 06-OCT-2003
DEFINITION Homo sapiens hypothetical protein MGC17791, mRNA (cdna clone
MGC:17791 IMAGE:388599), complete cds.
ACCESSION BC017672
VERSION BC017672.1 GI:17389225
KEYWORDS MGC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 2209)
Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,

```

```

Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
Schetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
Pahey,J., Helton,E., Kettman,M., Madan,A., Rodrigues,S.,
Sanchez,A., Whitting,M., Madan,A., Young,A.C., Shevchenko,Y.,
Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
Butterfield,Y.S., Zzywinski,M.I., Skalska,U., Smallos,D.E.,
Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
22388257
12477932
2 (bases 1 to 2209)
Strausberg,R.
Direct Submission
Submitted (03-DEC-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgabbs-remail.nih.gov
Tissue Procurement: DCTD/DP/Gazdar
cdna Library Preparation: Life Technologies, Inc.
cdna Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: http://www-shgc.stanford.edu
Contact: (Dickson, Mark) mcd@paxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.

```

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IPAX Plate: 22 Row: C Column: 16.

FEATURES

```

source
1..2209
  /organism="Homo sapiens"
  /mol_type="mRNA"
  /db_xref="taxon:9606"
  /clone_image="MGC:17791 IMAGE:388599"
  /tissue_type="Lung, carcinoma, large cell
  undifferentiated"
  /clone_lib="NIH MGC_69"
  /lab_host="DH10B"
  /note="vector: pCMV-SPORT6"
  1..2209
    /gene="MGC17791"
    /db_xref="LocusID:126282"
    116..676
      /codon_start=1
      /product="hypothetical protein MGC17791"
      /protein_id="AAH17672.1"
      /db_xref="GI:17389226"
      /db_xref="LocusID:126282"
      /translation="MDTFSTKSLAQKLLSKMASKAVVAVLDDTSSEVLDELVR
      ATREFTSRKEAOKMLNVLKVALKGLLQDGLGGEFLALLRRFRHRCCLAMTAV
      SFHQYDFTFDRVLAVGLLECRDLLHQAVGHILTAKSHGRINHVFGLHDCDFLAALY
      GPAPYRSHLRRICEGJGRMLDEGSL"

```

gene

CDS

ORIGIN

```

Query Match      8.7%; Score 85.8; DB 9; Length 2209;
Best Local Similarity 85.0%; Pred. No. 4.7e-10;
Matches 96; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

QY 34 CGATCTCTGACCTGCTGATCGCGCGCCCTCAGCTTCCCAAGTCTGGGATTACAGGCA 93

```



```

/rpt_family="AluSc"
complement(9688..9707)
/note="Single clone coverage"
9773..9801
/rpt_family="(TA)n"
complement(9835..10132)
/rpt_family="AluSc"
complement(9883..9971)
/note="single clone coverage"
complement(9915..9931)
/note="x30 qual single clone coverage"
complement(9943..9947)
/note="x30 qual single clone coverage"
10162..10771
/rpt_family="L1ME"
10781..11096
/rpt_family="AluSc"
complement(11118..11474)
/rpt_family="LTR47A"
11475..11796
/rpt_family="AluSc"
complement(11797..11833)
/rpt_family="LTR47A"
11839..12030
/rpt_family="L1ME"
12068..12130
/rpt_family="L1ME"
12131..12420
/rpt_family="MER89"
complement(12984..13280)
/rpt_family="AluSc"
complement(13351..13523)
/rpt_family="AluJo"
complement(13524..13831)
/rpt_family="AluSc"
complement(13832..13955)
/rpt_family="AluJo"
14482..14815
/rpt_family="AluSc"
14944..15132
/rpt_family="AluJo"
complement(15166..15605)

Query Match      8.7%; Score 85.8; DB 9; Length 185378;
Best Local Similarity 80.1%; Pred.No. 5e-10;
Matches 113; Conservative 0; Mismatches 27; Indels 1; Gaps 1;

Qy 34 CGATCTCTGACCTCGTGATCCGCCCGCTCAGCTTCCCAAGTCTGGGATTACAGGCA 93
Db 142821 CGATCTCTGACCTCGTGATCCGCCCGCTCAGCTTCCCAAGTCTGGGATTACAGGCG 142880

Qy 94 TGAGCCACCGCGCGCGCGCGCGCGCGACCTAGTTTAAAGGCGCCCTGCTGTTCTGCCCG 153
Db 142881 TGAGCCACCGCGCGCGCGCGCGCGCGACACACACACACACACACACACACACAC 142939

Qy 154 TGCGCGCGCTCCGAGCTGCC 174
Db 142940 TTCTCTGGCTCCCTACTGCC 142960

RESULT 14
AC022285
LOCUS AC022285 190466 bp DNA linear HTG 24-JAN-2002
DEFINITION Homo sapiens chromosome 17 clone RP11-14J13, WORKING DRAFT
SEQUENCE 8 unordered pieces.
AC022285
VERSION AC022285.41 GI:18308288
KEYWORDS HTG; HTGS PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 190466)

```

AUTHORS

Abola, A.P., Bruno, D., Conn, L., Della Rosa, M., Faulkner, D.,
 Federspiel, N., Glukhov, S., Hansen, N., Herman, Z.S., Hyman, R.,
 Mao, J., Komp, C., Kottler, S., Lam, B., Marathe, R., Miranda, M.,
 Morehouse, A.J., Nguyen, M., Oefner, P., Palm, C.J., Ramirez, D.,
 Southwick, A.M., Webb, C., Wilhelmy, J., Yu, S. and Davis, R.W.

JOURNAL

2 (bases 1 to 190466)

REFERENCE

Bruno, D., Conn, L., Della Rosa, M., Federspiel, N., Foreman, P.,
 Glukhov, S., Hansen, N., Hyman, R., Mao, J., Marathe, R.,
 Morehouse, A.J., Oefner, P., Palm, C.J., Ramirez, D., Wilhelmy, J.,
 Yu, S. and Davis, R.W.

TITLE

Direct Submission

JOURNAL

Submitted (28-JAN-2000) DNA Sequencing and Technology Center,
 Stanford University, 855 California Avenue, Palo Alto, CA 94304,
 USA

COMMENT

On Jan 24, 2002 this sequence version replaced gi:16903165.

----- Genome Center

Center: Stanford DNA Sequencing and Technology Development

Center code: SDSTDC

Web site: <http://sequence-www.stanford.edu/group/human/>

Contact: hum-info@sequence.stanford.edu

----- Project Information

Center project name: 760

Center clone name: RP11-14J13

----- Summary Statistics

Sequencing Vector: M13mp18; X02513; 98% of reads

Chemistry: Dye-primer; 16% of reads

Assembly program: Phrap; version 0.990319

Consensus quality: 185754 bases at least Q40

Consensus quality: 186781 bases at least Q30

Consensus quality: 187406 bases at least Q20

Insert size: 184236; agarose-fp

Quality coverage: 13.0x in Q20 bases; agarose-fp

Quality coverage: 12.6x in Q20 bases; sum-of-contigs.

* NOTE: This is a 'working draft' sequence. It currently

* is not known and their order in this sequence record is

* arbitrary. Gaps between the contigs are represented as

* runs of N, but the exact sizes of the gaps are unknown.

* This record will be updated with the finished sequence

* as soon as it is available and the accession number will

* be preserved.

* 1 1958: contig of 1958 bp in length

* 1959 2058: gap of unknown length

* 2059 4896: contig of 2838 bp in length

* 4897 4996: gap of unknown length

* 4997 8678: contig of 3682 bp in length

* 8679 8778: gap of unknown length

* 8779 11477: contig of 2699 bp in length

* 11478 11577: gap of unknown length

* 11578 18561: contig of 4984 bp in length

* 18562 16661: gap of unknown length

* 16662 39095: contig of 22434 bp in length

* 39096 39195: gap of unknown length

* 39196 77173: contig of 37978 bp in length

* 77174 77273: gap of unknown length

* 77274 190466: contig of 113193 bp in length.

FEATURES

source

1..190466

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/chromosome="17"

/clone="RP11-14J13"

/clone_lib="RP11 human BAC library 11"

1..1958

/note="assembly_name:Contig37"

2059..4896

/note="assembly_name:Contig38"

misc_feature

misc_feature

```

misc_feature 4997..8678
              /note="assembly_name:Contig39"
misc_feature 8779..11477
              /note="assembly_name:Contig40
clone_end:SP6"
              clone_end:16561
misc_feature 11578..16561
              /note="assembly_name:Contig41"
misc_feature 16662..39095
              /note="assembly_name:Contig42"
misc_feature 39196..77173
              /note="assembly_name:Contig43"
misc_feature 77274..190466
              /note="assembly_name:Contig44
clone_end:17"

ORIGIN
Query Match      8.7%; Score 85.8; DB 2; Length 190466;
Best Local Similarity 80.1%; Pred. No. 5e-10;
Matches 113; Conservative 0; Mismatches 27; Indels 1; Gaps 1;

Qy 34 CGATCTCTGACCTCGTGATCGCCGCCCTCAGCTCCCAAAGTGTGGGATTACAGGCA 93
      |||||
Db 147904 CGATCTCTGACCTCGTGATCGCCGCCCTCAGCTCCCAAAGTGTGGGATTACAGGCG 147963

Qy 94 TGAGCCACCGCGCCGCCGCCGACACCTAGTTTAAAGGCGCCCTGCTTCTGCTGCCGC 153
      |||||
Db 147964 TGAGCCACCGCGCCGCCGCCGAGACACACATTTTACAGGACTGCCCTGTTCCAAACC 148022

Qy 154 TGCGCGCGCTCCCGACGTCGCC 174
      |||||
Db 148023 TTCTCTGGCTCCCTACTGCC 148043

```

```

RESULT 15
AC129497
LOCUS AC129497 199463 bp DNA linear HTG 27-MAR-2003
DEFINITION Homo sapiens chromosome 17 clone RP11-945P9 map 17, 9 unordered
          pieces.
ACCESSION AC129497
VERSION AC129497.3 GI:22758749
KEYWORDS HTG; HTGS PHASE1; HTGS_FULLTOP; HTGS_CANCELLED.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 199463)
AUTHORS Birren,B., Nusbaum,C. and Lander,E.
TITLE Homo sapiens chromosome 17, clone RP11-945P9
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 199463)
AUTHORS Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
          Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L., Boukhgalter,B.,
          Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,
          Cook,A., Cooke,P., DeArelano,K., Dewar,K., Diaz,J.S., Dodge,S.,
          Faro,S., Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J.,
          Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B.,
          Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A.,
          Karatas,A., Kellis,C., Landers,T., Levine,R., Lindblad-Toh,K.,
          Liu,G., Maclean,C., Macdonald,P., Major,J., Matthews,C.,
          McCarthy,M., Meldrum,J., Meneus,L., Mihova,T., Mienga,V.,
          Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H.,
          O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K.,
          Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P.,
          Roman,J., Roy,A., Schauer,S., Schupback,R., Seaman,S., Severy,P.,
          Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J.,
          Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H.,
          Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J.,
          Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (30-JUL-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE 3 (bases 1 to 199463)
AUTHORS Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,

```

```

Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L., Boukhgalter,B.,
Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,
Cook,A., Cooke,P., DeArelano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Faro,S., Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J.,
Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B.,
Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A.,
Karatas,A., Kellis,C., Landers,T., Levine,R., Lindblad-Toh,K.,
Liu,G., Maclean,C., Macdonald,P., Major,J., Matthews,C.,
McCarthy,M., Meldrum,J., Meneus,L., Mihova,T., Mienga,V.,
Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H.,
O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K.,
Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P.,
Roman,J., Roy,A., Schauer,S., Schupback,R., Seaman,S., Severy,P.,
Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J.,
Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J.,
Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (27-MAR-2003) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Sep 8, 2002 this sequence version replaced gi:22475392.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
-----
Center: Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: W18R
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
-----
Project Information
Center project name: L27853
Center clone name: 945_P_9
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 9 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 1451: contig of 1451 bp in length
* 1452 1551: gap of 100 bp
* 1552 5314: contig of 3763 bp in length
* 5315 5414: gap of 100 bp
* 5415 31286: contig of 25872 bp in length
* 31287 31386: gap of 100 bp
* 31387 60741: contig of 29355 bp in length
* 60742 60841: gap of 100 bp
* 60842 159763: contig of 98922 bp in length
* 159764 159863: gap of 100 bp
* 159864 186613: contig of 26750 bp in length
* 186614 186713: gap of 100 bp
* 186714 190447: contig of 3734 bp in length
* 190448 190547: gap of 100 bp
* 190548 195762: contig of 5215 bp in length
* 195763 195862: gap of 100 bp
* 195863 199463: contig of 3601 bp in length.
*
* Location/Qualifiers
* 1..199463
* /organism="Homo sapiens"
* /mol_type="genomic DNA"
* /db_xref="taxon:9606"
* /chromosomes="17"
* /map="17"
* /clone_lib="RP11-945P9"
* /clone_lib="RP11-945P9"

```

ORIGIN

```

Query Match      8.7%; Score 85.8; DB 2; Length 199463;
Best Local Similarity 80.1%; Pred. No. 5e-10;
Matches 113; Conservative 0; Mismatches 27; Indels 1; Gaps 1;

```


CC tamoxifen resistant, or to treat osteoporosis, by increasing levels of
 CC ERCA3 in cells. The encoding polynucleotide can be used to inhibit
 CC translation of a mRNA encoding ERCA3. ERCA3 acts as a coregulator
 CC protein and can bind to the estrogen receptor to activate a molecular
 CC cellular response in the cell. The present sequence represents the human
 CC ERCA3 encoding cDNA
 XX
 SQ Sequence 990 BP; 189 A; 281 C; 317 G; 203 T; 0 U; 0 Other;

Query Match 100.0%; Score 990; DB 6; Length 990;

Best Local Similarity 100.0%; Pred. No. 3,7e-235;

Matches 990; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGAATTGTTCTGAGGCCAAGAAATTCGCAAGGCGATCTCCTGACCTCGTGTATCCGCCCG 60
 Db 1 GGAATTGTTCTGAGGCCAAGAAATTCGCAAGGCGATCTCCTGACCTCGTGTATCCGCCCG 60
 QY 61 CCTCAGCTTCCCAAGTCTGGGATTACAGGCATGAGCCACCGCGCCCGCCCGCCGACAC 120
 Db 61 CCTCAGCTTCCCAAGTCTGGGATTACAGGCATGAGCCACCGCGCCCGCCCGCCGACAC 120
 QY 121 CTAGTTTAAAGGGCCCTGCTGTGCTGCGCTGCGCGCTCCAGCTGCCAGTCTG 180
 Db 121 CTAGTTTAAAGGGCCCTGCTGTGCTGCGCTGCGCGCTCCAGCTGCCAGTCTG 180
 QY 181 GCGGCTCAGTCCCGCGTTCATGTGTGGAGACCGCTCGCTGAGCGCTGATGTGG 240
 Db 181 GCGGCTCAGTCCCGCGTTCATGTGTGGAGACCGCTCGCTGAGCGCTGATGTGG 240
 QY 241 CTTGCTGATGCAATTTGACCCGGCTCTGACCTGGGCTAGGGGAAAGGCGAGGCGCG 300
 Db 241 CTTGCTGATGCAATTTGACCCGGCTCTGACCTGGGCTAGGGGAAAGGCGAGGCGCG 300
 QY 301 AATTGGGCGCGAGGGCCAGGCTCGCGACCCCGCTGCGCTCCCGTGGCGCCCGCAG 360
 Db 301 AATTGGGCGCGAGGGCCAGGCTCGCGACCCCGCTGCGCTCCCGTGGCGCCCGCAG 360
 QY 361 CGCTCCCGTGGCGCTGGAGTGAGTCTTACCGTCCGAGATCGTCCGCAACTGGCGCA 420
 Db 361 CGCTCCCGTGGCGCTGGAGTGAGTCTTACCGTCCGAGATCGTCCGCAACTGGCGCA 420
 QY 421 GCTGTGATGGGGCTGCTGAGCGCTGCTGTTGTTGATTTGATTTGATTTGATTTG 480
 Db 421 GCTGTGATGGGGCTGCTGAGCGCTGCTGTTGTTGATTTGATTTGATTTGATTTG 480
 QY 481 TGTGTCTCTGAAGACATGGACATAGTCTGGAGGGTCTTGAAGAGTGTATCCCGGCC 540
 Db 481 TGTGTCTCTGAAGACATGGACATAGTCTGGAGGGTCTTGAAGAGTGTATCCCGGCC 540
 QY 541 CACATCAATGGGCTTAGTCTAGAACGGGTGTGGTGGGCTTAGGGCGAGGGCG 600
 Db 541 CACATCAATGGGCTTAGTCTAGAACGGGTGTGGTGGGCTTAGGGCGAGGGCG 600
 QY 601 CAGACATACCCGGAAGTGTGATTTATACCGCAAGGGCTGGATGAAACCCCGCAA 660
 Db 601 CAGACATACCCGGAAGTGTGATTTATACCGCAAGGGCTGGATGAAACCCCGCAA 660
 QY 661 GACACTGGAAGCTGTGTGGCTGAGAGGGCCCGCAATCCAGTGTGTGGGCTTTAC 720
 Db 661 GACACTGGAAGCTGTGTGGCTGAGAGGGCCCGCAATCCAGTGTGTGGGCTTTAC 720
 QY 721 AGGAAGAGCTCCACCTCTTGGAGTGTGAGATGGATAGTGTGTCACCCGATGG 780
 Db 721 AGGAAGAGCTCCACCTCTTGGAGTGTGAGATGGATAGTGTGTCACCCGATGG 780
 QY 781 GAGCTGGGGCGGGCAGATGTGCCCCAGTCAAAAGCTGATTTGACCTGGGGCTCTG 840
 Db 781 GAGCTGGGGCGGGCAGATGTGCCCCAGTCAAAAGCTGATTTGACCTGGGGCTCTG 840
 QY 841 GACTTCCCTGATTTCTGCTGATCTCCAGCAAGTCTGTCCTGGTGGCTTCAT 900
 Db 841 GACTTCCCTGATTTCTGCTGATCTCCAGCAAGTCTGTCCTGGTGGCTTCAT 900

QY 901 CCACTCTCTCACTTCTCTGCTTCAAGTAAATTCAGAGATCTGTGGTGCAGAAAAA 960
 Db 901 CCACTCTCTCACTTCTCTGCTTCAAGTAAATTCAGAGATCTGTGGTGCAGAAAAA 960
 QY 961 AAAAAAATGAGCGCGCGCAAGTT 990
 Db 961 AAAAAAATGAGCGCGCGCAAGTT 990

RESULT 2

AAS69040/c

ID AAS69040 standard; cDNA; 597 BP.

XX AAS69040;

DT 13-FEB-2002 (first entry)

XX DNA encoding novel human diagnostic protein #4844.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder; ss.

XX Homo sapiens.

XX WO200175067-A2.

XX 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US008631.

XX 31-MAR-2000; 2000US-00540217.

XX 23-AUG-2000; 2000US-00649167.

XX (HYSE-) HYSEQ INC.

XX Drmanac RT, Liu C, Tang YT;

XX WPI; 2001-639362/73.

XX P-PSDB; ABG04853.

PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity.

XX Claim 1; SEQ ID NO 4844; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
 CC sequences. (I) is useful as hybridisation probes, polymerase chain
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
 CC and in recombinant production of (II). The polynucleotides are also used
 CC in diagnostics as expressed sequence tags for identifying expressed
 CC genes. (I) is useful in gene therapy techniques to restore normal
 CC activity of (II) or to treat disease states involving (II). (II) is
 CC useful for generating antibodies against it, detecting or quantitating a
 CC polypeptide in tissue, as molecular weight markers and as a food
 CC supplement. (II) and its binding partners are useful in medical imaging
 CC of sites expressing (II). (I) and (II) are useful for treating disorders
 CC involving aberrant protein expression or biological activities. The
 CC polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
 CC coding sequences of the invention. Note: The sequence data for this
 CC patent did not appear in the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 597 BP; 97 A; 217 C; 196 G; 87 T; 0 U; 0 Other;

Query Match

Best Local Similarity 15.0%; Score 148.6; DB 5; Length 597;

Pred. No. 6.9e-27;

Matches 157; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
QY 128 TAAAGGCGCCCTGCTGTTGCTGCGCTGCGCGCGCTCCAGCTGCCAGTCTGGCGGCT 187
Db 597 TTAGAGGCCCTGCTGTTGCTGCGCTGCGCGCTCCAGCTGCCAGTCTGGCGGCT 538
QY 188 CAGTCCCGCTTGGCATGTGTGGAGACCGCTGCGGTAAAGCCTGGATGTGGCTTCGCT 247
Db 537 CAGTCCCGCTTGGCATGTGTGGAGACCGCTGCGGTAAAGCCTGGATGTGGCTTCGCT 478
QY 248 GATGCACATTGGACCGGCTCTGACTGGGCTAGGGAAGGGCAGGAGGCG 298
Db 477 GATGCACATTGGACCGGCTCTGACTGGGCTAGGGAAGGGCAGGAGGCG 427
RESULT 3
AAH18715/c
ID AAH18715 standard; cDNA; 1602 BP.
XX
AC AAH18715;
XX
DT 26-JUN-2001 (first entry)
XX
DE Human cDNA sequence SEQ ID NO:18986.
XX
KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX
OS Homo sapiens.
XX
FN EP1074617-A2.
XX
PD 07-FEB-2001.
XX
PF 28-JUL-2000; 2000EP-00116126.
XX
PR 29-JUL-1999; 99JP-00248036.
PR 27-AUG-1999; 99JP-00300253.
PR 11-JAN-2000; 2000JP-00118776.
PR 02-MAY-2000; 2000JP-00183767.
PR 09-JUN-2000; 2000JP-00241899.
XX
PA (HELI-) HELIX RES INST.
XX
PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX
DR WPI; 2001-318749/34.
XX
PT Primer sets for synthesizing polynucleotides, particularly the 5602 full-length cDNAs defined in the specification, and for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs.
XX
PS Claim 8; SEQ ID NO 18986; 2537bp + Sequence Listing; English.
XX
CC The present invention describes primer sets for synthesizing 5602 full-length cDNAs defined in the specification. Where a primer set comprises:
CC (a) an oligo-dr primer and an oligonucleotide complementary to the
CC complementary strand of a polynucleotide which comprises one of the 5602
CC nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in the
CC specification. The primer sets can be used in antisense therapy and in
CC gene therapy. The primers are useful for synthesizing polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and

CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893
CC represent human amino acid sequences; and AAH13629 to AAH13632 represent
CC oligonucleotides, all of which are used in the exemplification of the
CC present invention
XX
SQ Sequence 1602 BP; 334 A; 458 C; 491 G; 319 T; 0 U; 0 Other;
Query Match 8.8%; Score 87.4; DB 4; Length 1602;
Best Local Similarity 85.8%; Pred. No. 1.4e-11;
Matches 97; Conservative 0; Mismatches 16; Indels 0; Gaps 0;
QY 34 CGATCTCTGACCTCGTGTGATCGCGCGCTCAGCTTCCCAAAGTCTGGATTACAGGCA 93
Db 935 CGATCTCTGACCTCGTGTGATCGCGCGCTCAGCTTCCCAAAGTCTGGATTACAGGCG 876
QY 94 TGAGCCACCGCGCGCGCGCGCGCACCTAGTCTTAAAGGCGCCCTGCTGTG 146
Db 875 TGAGCCACCG 823
RESULT 4
ADC85452
ID ADC85452 standard; DNA; 67782 BP.
XX
AC ADC85452;
XX
DT 01-JAN-2004 (first entry)
XX
DE Human Git2 genomic sequence.
XX
KW Cytostatic; gene therapy; vaccine; cancer; carcinoma-associated gene; CA;
KW secreted; transmembrane; intracellular; ds.
XX
OS Homo sapiens.
XX
PN WO2003045230-A2.
XX
PD 05-JUN-2003.
XX
PF 02-DEC-2002; 2002WO-US038582.
XX
PR 30-NOV-2001; 2001US-00997722.
XX
PA (SAGR-) SAGRES DISCOVERY.
XX
PI Morris DW, Engelhard EK;
XX
DR WPI; 2003-513603/48.
XX
PT New recombinant nucleic acid comprising a nucleotide sequence of any of
PT the carcinoma-associated (CA) genes, useful for screening for drug
PT candidates for diagnosing or treating carcinomas.
XX
PS Claim 1; SEQ ID NO 238; 983pp; English.
XX
CC The invention relates to a recombinant nucleic acid comprising a
CC nucleotide sequence selected from any of the fully defined carcinoma-
CC associated (CA) genes from the 50 tables given in the specification. The
CC CA proteins are secreted, transmembrane or intracellular proteins. The
CC recombinant nucleic acids are useful for screening for drug candidates
CC for diagnosing or treating carcinomas. Sequences given in ADC85215-
CC ADC85514 represent CA genes of the invention.
XX
SQ Sequence 67782 BP; 18941 A; 14566 C; 14980 G; 18926 T; 0 U; 369 Other;
Query Match 8.5%; Score 84.2; DB 9; Length 67782;
Best Local Similarity 81.0%; Pred. No. 3e-10;
Matches 98; Conservative 0; Mismatches 23; Indels 0; Gaps 0;
QY 34 CGATCTCTGACCTCGTGTGATCGCGCGCTCAGCTTCCCAAAGTCTGGATTACAGGCA 93
Db 51741 CGATCTCTGACCTCGTGTGATCGCGCGCTCAGCTTCCCAAAGTCTGGATTACAGGCG 51800

Qy 154 T 154
Db 51862 T 51862

RESULT 7
AAC19185
ID AAC19185 standard; cDNA; 274 BP.
XX AC AAC19185;
XX DT 06-OCT-2000 (first entry)
XX DE Human secreted protein 5' EST, SEQ ID NO: 23260.
XX KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KW Gene therapy; chromosome mapping; ss.
XX OS Homo sapiens.
XX PN EPI033401-A2.
XX PD 06-SEP-2000.
XX PF 21-FEB-2000; 2000EP-00200610.
XX PR 26-FEB-1999; 99US-0122487P.
XX PA (GEST) GENSET.
XX PI Dumas Milne Edwards J, Duclert A, Giordano J;
XX WPI; 2000-500381/45.
XX DR New nucleic acid that is a 5' expressed sequence tag (5' EST) for
XX PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
XX PT diagnostic, forensic, gene therapy and chromosome mapping procedures.
XX PS Claim 1; SEQ ID NO 23260; 71pp + Sequence Listing; English.
XX CC The present sequence is one of a large number of 5' ESTs derived from
CC mRNAs encoding secreted proteins. No ORF has yet been conclusively
CC identified within the present sequence. The 5' ESTs were prepared from
CC total human RNAs or polyA+ RNAs derived from 30 different tissues. EST
CC sequences usually correspond mainly to the 3' untranslated region (UTR)
CC of the mRNA because they are often obtained from oligo-dT primed cDNA
CC libraries. Such ESTs are not well suited for isolating cDNA sequences
CC derived from the 5' ends of mRNAs and even in those cases where longer
CC cDNA sequences have been obtained, the full 5' UTR is rarely included. 5'
CC ESTs are derived from mRNAs with intact 5' ends and can therefore be used
CC to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used in
CC diagnostic, forensic, gene therapy and chromosome mapping procedures.
CC They are used to obtain upstream regulatory sequences and to design
CC expression and secretion vectors
XX SQ Sequence 274 BP; 50 A; 88 C; 59 G; 77 T; 0 U; 0 Other;

Query Match 8.4%; Score 83.6; DB 3; Length 274;
Best Local Similarity 90.8%; Pred. No. 6.7e-11;
Matches 89; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 34 CGATCTCTGACCTCGTGATCGGCCCGCCTCAGTTCCTCCAAAGTCGGGATTACAGGCA 93
Db 126 CGATCTCTCCTCCTCGTGATCGGCCCGCCTCGGCTCCCAAGTCGTGGGATTACAGGCA 185
Qy 94 TGAGCCACCGCGCGCGCGCGCCGACACTAGTTTAA 131
Db 186 TGAGCCACCGCGCGCGCGCCCAATCATCTGTTTAA 223

RESULT 8
AAH99495/c
ID AAH99495 standard; cDNA; 2105 BP.

XX AC AAH99495;
XX DT 16-OCT-2001 (first entry)
XX DE Human protein encoding cDNA sequence SEQ ID NO:330.
XX KW Human; cancer; ulcer; HIV infection; human immunodeficiency virus;
KW antiinflammatory; antirheumatic; antiarthritic; immunosuppressive;
KW antibacterial; endocrine; cardiant; central nervous system; virucide;
KW anti-HIV; fungicide; antimutagen; cardiovascular; antianaemic; anaemia;
KW antiagregant; haemostatic; vulnary; antileuc; osteopathic; eczema;
KW dermatological; antiallergic; antiasthmatic; antidiabetic; cytostatic;
KW neuroprotective; antidepressant; nootropic; antiparkinsonian; infection;
KW immunostimulant; gene therapy; antisense therapy; vaccine; inflammation;
KW antianaphylactic; rheumatoid arthritis; septic shock; pancreatitis;
KW cardiac dysfunction; neuropathology; cardiac anaphylaxis; autoimmunity;
KW genetic disease; haematopoietic disorder; platelet disorder; asthma;
KW thrombocytopaenia; osteoporosis; severe combined immunodeficiency;
KW allergic rhinitis; diabetes; multiple sclerosis; depression;
KW Alzheimer's disease; Parkinson's disease; neurodegenerative disorder;
KW neurological disorder; ss.
XX OS Homo sapiens.
XX PN WO200153455-A2.
XX PD 26-JUL-2001.
XX PF 22-DEC-2000; 2000WO-US035017.
XX PR 23-DEC-1999; 99US-00471275.
XX PR 21-JAN-2000; 2000US-00488725.
XX PR 25-APR-2000; 2000US-00552317.
XX PA (HYSE-) HYSEQ INC.
XX PI Tang YT, Liu C, Drmanac RT;
XX WPI; 2001-457603/49.
XX P-PSDB; AAH25554.
XX DR Isolated human polynucleotides encoding polypeptides, useful for the
XX PT treatment and diagnosis of e.g. cancer, ulcers and HIV infection.
XX PS Claim 1; Page 444; 1217pp; English.
XX CC AAH99166 to AAH99904 encode the human proteins given in AAH25225 to
CC AAH25963. The proteins can have activities based on the tissues and cells
CC they are expressed in, such as: antiinflammatory; antirheumatic;
CC antiarthritic; immunosuppressive; antibacterial; endocrine; cardiant;
CC central nervous system; virucide; anti-HIV; fungicide; antimutagen;
CC cardiovascular; antianaemic; antiagregant; haemostatic; vulnary;
CC antileuc; osteopathic; dermatological; antiallergic; antiasthmatic;
CC antidiabetic; cytostatic; neuroprotective; antidepressant; nootropic;
CC antiparkinsonian; and immunostimulant. The proteins and polynucleotides
CC encoding them can be used in gene therapy, antisense therapy and vaccine
CC production. The proteins and polynucleotides are useful for screening for
CC agonists or antagonists of a protein and for the treatment and diagnosis
CC of disorders associated with the activity of a protein e.g. inflammation,
CC rheumatoid arthritis, septic shock, pancreatitis, cardiac dysfunction,
CC neuropathology, cardiac anaphylaxis, viral, bacterial, HIV and fungal
CC infections, autoimmunity, genetic diseases, haematopoietic disorders,
CC anaemia, platelet disorders, thrombocytopaenia, wounds, burns, ulcers,
CC osteoporosis, severe combined immunodeficiency, eczema, allergic
CC rhinitis, asthma, diabetes, cancer, multiple sclerosis, depression,
CC Alzheimer's disease, Parkinson's disease, neurodegenerative and
CC neurological disorders
XX SQ Sequence 2105 BP; 580 A; 460 C; 582 G; 483 T; 0 U; 0 Other;

Query Match 8.4%; Score 83.6; DB 4; Length 2105;
Best Local Similarity 90.8%; Pred. No. 1.3e-10;

PR 17-NOV-2000; 2000US-0249217P.
 PR 17-NOV-2000; 2000US-0249218P.
 PR 17-NOV-2000; 2000US-0249244P.
 PR 17-NOV-2000; 2000US-0249245P.
 PR 17-NOV-2000; 2000US-0249264P.
 PR 17-NOV-2000; 2000US-0249265P.
 PR 17-NOV-2000; 2000US-0249297P.
 PR 17-NOV-2000; 2000US-0249299P.
 PR 17-NOV-2000; 2000US-0249300P.
 PR 01-DEC-2000; 2000US-0250160P.
 PR 01-DEC-2000; 2000US-0250391P.
 PR 05-DEC-2000; 2000US-0251030P.
 PR 05-DEC-2000; 2000US-0251988P.
 PR 05-DEC-2000; 2000US-0256719P.
 PR 06-DEC-2000; 2000US-0251479P.
 PR 08-DEC-2000; 2000US-0251856P.
 PR 08-DEC-2000; 2000US-0251868P.
 PR 08-DEC-2000; 2000US-0251869P.
 PR 08-DEC-2000; 2000US-0251989P.
 PR 11-DEC-2000; 2000US-0251990P.
 PR 05-JAN-2001; 2001US-0259678P.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Barash SC, Ruben SM;
 XX
 PI WPI; 2001-483426/52.
 DR
 XX
 PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
 PT useful for preventing, diagnosing and/or treating cancers and metastasis.
 XX
 PS Disclosure; SEQ ID NO 21235; 3071pp + Sequence Listing; English.
 XX
 CC AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
 CC amino acid sequences given in AAK82170 to AAK91921. (I) have cytostatic
 CC activity, and can be used in gene therapy and vaccine production. (I)
 CC proteins and polynucleotides may be used in the prevention, diagnosis and
 CC treatment of diseases associated with inappropriate (I) expression. For
 CC example, they may be used to treat disorders associated with decreased
 CC expression by rectifying mutations or deletions in a patient's genome
 CC that affect the activity of (I) by expressing inactive proteins or to
 CC supplement the patients own production of (I). Additionally, (I)
 CC polynucleotides may be used to produce the secreted (I), by inserting the
 CC nucleic acids into a host cell and culturing the cell to express the
 CC protein. (I) proteins and polynucleotides may be used to prevent,
 CC diagnose and treat immune/haematopoietic-related diseases, especially
 CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
 CC to AAK87694 represent human immune/haematopoietic antigen genomic
 CC sequences from the present invention. AAK54942 to AAK54950 and AAK82169
 CC represent sequences used in the exemplification of the present invention
 XX
 SQ Sequence 2214 BP; 617 A; 455 C; 523 T; 0 U; 0 Other;
 Query Match 8.4%; Score 83.2; DB 4; Length 2214;
 Best Local Similarity 83.9%; Pred. No. 1.7e-10;
 Matches 94; Conservative 0; Mismatches 18; Indels 0; Gaps 0;
 QY 34 CGATCTCTGACCTCGTGACCGCCGCCCTCCAGCTCCCAAGTCTGGGATTACAGGCA 93
 Db 403 CGATCTCTGACCTCGTGACCGCCGCCCTCCAGCTCCCAAGTCTGGGACTGCGAGG 344
 QY 94 TGAGCCACCGCGCCGCCCGCCGACACCTAGTTTAAAGGGCCCCCTGCTGT 145
 Db 343 TGAGCCACCGCGCCGCCCGCCGACGTCATTTCTTTTCAGGAGTCCCTTCTGTT 292
 RESULT 10
 AAK66424/c
 ID AAK66424 standard; DNA; 6742 BP.
 XX
 AC AAK66424;
 XX

DT 06-NOV-2001 (first entry)
 XX Human immune/haematopoietic antigen genomic sequence SEQ ID NO:21236.
 DE Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
 XX cytostatic; gene therapy; vaccine; metastasis; ds.
 KW Homo sapiens.
 XX WO200157182-A2.
 XX 09-AUG-2001.
 XX
 XX 17-JAN-2001; 2001WO-US001354.
 XX 31-JAN-2000; 2000US-0179065P.
 PR 04-FEB-2000; 2000US-0180628P.
 PR 24-FEB-2000; 2000US-0184664P.
 PR 02-MAR-2000; 2000US-0186350P.
 PR 16-MAR-2000; 2000US-0189874P.
 PR 17-MAR-2000; 2000US-0190076P.
 PR 18-APR-2000; 2000US-0198123P.
 PR 19-MAY-2000; 2000US-0205515P.
 PR 07-JUN-2000; 2000US-0209467P.
 PR 28-JUN-2000; 2000US-0214886P.
 PR 30-JUN-2000; 2000US-0215135P.
 PR 07-JUL-2000; 2000US-0216647P.
 PR 07-JUL-2000; 2000US-0216880P.
 PR 11-JUL-2000; 2000US-0217487P.
 PR 11-JUL-2000; 2000US-0217496P.
 PR 14-JUL-2000; 2000US-0218290P.
 PR 26-JUL-2000; 2000US-0220963P.
 PR 26-JUL-2000; 2000US-0220964P.
 PR 14-AUG-2000; 2000US-0224518P.
 PR 14-AUG-2000; 2000US-0224519P.
 PR 14-AUG-2000; 2000US-0225213P.
 PR 14-AUG-2000; 2000US-0225214P.
 PR 14-AUG-2000; 2000US-0225266P.
 PR 14-AUG-2000; 2000US-0225267P.
 PR 14-AUG-2000; 2000US-0225268P.
 PR 14-AUG-2000; 2000US-0225270P.
 PR 14-AUG-2000; 2000US-0225447P.
 PR 14-AUG-2000; 2000US-0225757P.
 PR 14-AUG-2000; 2000US-0225758P.
 PR 14-AUG-2000; 2000US-0225759P.
 PR 18-AUG-2000; 2000US-0226279P.
 PR 22-AUG-2000; 2000US-0226681P.
 PR 22-AUG-2000; 2000US-0226686P.
 PR 23-AUG-2000; 2000US-0227182P.
 PR 30-AUG-2000; 2000US-0227009P.
 PR 01-SEP-2000; 2000US-0228924P.
 PR 01-SEP-2000; 2000US-0228287P.
 PR 01-SEP-2000; 2000US-0229343P.
 PR 01-SEP-2000; 2000US-0229344P.
 PR 05-SEP-2000; 2000US-0229345P.
 PR 05-SEP-2000; 2000US-0229509P.
 PR 06-SEP-2000; 2000US-0229513P.
 PR 06-SEP-2000; 2000US-0230437P.
 PR 08-SEP-2000; 2000US-0230438P.
 PR 08-SEP-2000; 2000US-0231242P.
 PR 08-SEP-2000; 2000US-0231243P.
 PR 08-SEP-2000; 2000US-0231244P.
 PR 08-SEP-2000; 2000US-0231413P.
 PR 08-SEP-2000; 2000US-0231414P.
 PR 08-SEP-2000; 2000US-0232080P.
 PR 12-SEP-2000; 2000US-0232081P.
 PR 14-SEP-2000; 2000US-0232397P.
 PR 14-SEP-2000; 2000US-0232398P.
 PR 14-SEP-2000; 2000US-0232399P.
 PR 14-SEP-2000; 2000US-0232400P.
 PR 14-SEP-2000; 2000US-0232401P.
 PR 14-SEP-2000; 2000US-0233063P.

CC	AAK54951	to AAK64702 encode the human immune/haematopoietic antigen (I)
CC	amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic	
CC	activity, and can be used in gene therapy and vaccine production. (I)	
CC	proteins and polynucleotides may be used in the prevention, diagnosis and	
CC	treatment of diseases associated with inappropriate (I) expression. For	
CC	example, they may be used to treat disorders associated with decreased	
CC	expression by rectifying mutations or deletions in a patient's genome	
CC	that affect the activity of (I) by expressing inactive proteins or to	
CC	supplement the patients own production of (I). Additionally, (I)	
CC	polynucleotides may be used to produce the secreted (I), by inserting the	
CC	nucleic acids into a host cell and culturing the cell to express the	
CC	protein. (I) proteins and polynucleotides may be used to prevent,	
CC	diagnose and treat immune/haematopoietic-related diseases, especially	
CC	cancers and cancer metastases of haematopoietic-derived cells. AAK64703	
CC	to AAK87694 represent human immune/haematopoietic antigen genomic	
CC	sequences from the present invention. AAK54942 to AAK54950 and AAM82169	
CC	represent sequences used in the exemplification of the present invention	
XX		
SQ	Sequence 10772 BP; 3111 A; 2137 C; 2495 G; 3029 T; 0 U; 0 Other;	
	Query Match	8.4%; Score 83.2; DB 4; Length 10772;
	Best Local Similarity	83.9%; Pred. No. 2.9e-10;
	Matches	94; Conservative 0; Mismatches 18; Indels 0; Gaps 0;
QY	34	CGATCTCTGACCTGTGATCCGCCGCCTCAGCTTCCC AAAGTGCTGGGATTACAGCA 93
Dd	8950	CGATCTCTGACCTGTGATCCGCCGCCTCAGCTTCCC AAAGTGCTGGGACTGCAGCG 8891
QY	94	TGAGCCACC CGCCGCCGCCCGCACCTAGTTTTTA AAGGCCCTCTCTGTT 145
Dd	8890	TGAGCCACC CGCCGCCGCCCGCAACGTCA TTTTCTTTTCAGGAGTCCCTTCTGTT 8839
RESULT 12		
ID	ADC86452	standard; DNA; 13695 BP.
XX	AC	ADC86452;
XX	DT	01-JAN-2004 (first entry)
XX	Human	GPCR gene SEQ ID NO:905.
KW	ds; Gene; human; GPCR;	
XX	guanosine triphosphate-binding protein coupled receptor; gene therapy.	
OS	Homo sapiens.	
XX	EP1270724-A2.	
XX	02-JAN-2003.	
XX	18-JUN-2002; 2002EP-00013517.	
PR	18-JUN-2001; 2001JP-00246789.	
PA	(NAAD-) NAT INST ADVANCED IND SCI & TECHNOLOGY.	
PA	(ADSC-) CENT ADVANCED SCI & TECHNOLOGY INCUBATIO.	
PI	Suwa M, Asai K, Akiyama Y, Aburatani H;	
DR	WPI; 2003-315783/31.	
DR	P-PSDB; ADC86453.	
PT	New polynucleotide, useful for preparing a composition for treating a	
PT	patient in need of increased or suppressed activity or expression of the	
XX	guanosine triphosphate-binding protein coupled receptor.	
ES	Claim 1; SEQ ID NO 905; 28pp; English.	
CC	The invention relates to a novel polynucleotide encoding a guanosine	
CC	triphosphate-binding protein coupled receptor (GPCR). A polynucleotide of	
CC	the invention may have a use in gene therapy. The polynucleotide and	

CC polypeptide are useful for preparing a composition for treating a patient
CC in need of increased or suppressed activity or expression of the
CC guanine triphosphate-binding protein coupled receptor. The
CC polynucleotide sequences shown in AOC8548-ADC87616 encode GPCR's of the
XX invention.

XX Sequence 13695 BP; 2885 A; 3811 C; 3868 G; 3131 T; 0 U; 0 Other;
SQ

Query Match 8.4%; Score 83; DB 9; Length 13695;
Best Local Similarity 89.9%; Pred. No. 3.5e-10;
Matches 89; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 34 CGATCTCTGACCTCGTGATCCGCCCGCCTCAGCTTCCCCAAAGTCGTGGATTACAGGCA 93
Db 842 CGATCTCTGACCTCGTGATCCGCCCGCCTCAGCTTCCCCAAAGTCGTGGATTACAGGCG 901

Qy 94 TGAGCACCGCGCGCGCGCCCGCACACTAGTTTAAAG 132
 |||||
Db 902 TGAGCACCGCGCGCGCGCTCAGTTTCTCTTTTAG 940

RESULT 13
AAH17556/c
ID AAH17556 standard; cDNA; 2251 BP.
XX AC AAH17556;
XX XX
XX XX
DT 26-JUN-2001 (first entry)
XX XX
DE Human cDNA sequence SEQ ID NO:17037.
XX XX
KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX XX
OS Homo sapiens.
XX XX
FN EF1074617-A2.
XX XX
PD 07-FEB-2001.
XX XX
PF 28-JUL-2000; 2000EP-00116126.
XX XX
PR 29-JUL-1999; 99JP-00248036.
PR 27-AUG-1999; 99UP-00300253.
PR 11-JAN-2000; 2000JP-00118776.
PR 02-MAY-2000; 2000JP-00183767.
PR 09-JUN-2000; 2000JP-00241899.
XX XX
PA (HELI-) HELIX RES INST.
XX XX
PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX XX
WPI; 2001-318749/34.
DR
XX XX
PT Primer sets for synthesizing polynucleotides, particularly the 5602 full-length cDNAs defined in the specification, and for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs.
XX XX
PS Claim 8; SEQ ID NO 17037; 2537pp + Sequence Listing; English.
XX XX
CC The present invention describes primer sets for synthesising 5602 full-length cDNAs defined in the specification, where a primer set comprises:
CC (a) an oligo-dr primer and an oligonucleotide complementary to the
CC complementary strand of a polynucleotide which comprises one of the 5602
CC nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in the

CC specification. The primer sets can be used in antisense therapy and in
CC gene therapy. The primers are useful for synthesising polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to AAH95893
CC represent human amino acid sequences, and AAH13629 to AAH13632 represent
CC oligonucleotides, all of which are used in the exemplification of the
CC present invention

XX
SQ Sequence 2251 BP; 678 A; 479 C; 550 G; 544 T; 0 U; 0 Other;

Query Match 8.3%; Score 82.6; DB 4; Length 2251;
Best Local Similarity 73.1%; Pred. No. 2.4e-10;
Matches 106; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

QY 34 CGATCTCTGACCTCGTGTATCCGCCCGCTCAGCTTCCCAAGTGTGGGATTACAGCA 93
Db 1740 CGAACTCCAGACCTCGTGTATCCGCCCGCTCAGCTTCCCAAGTGTGGGATTACAGCA 1681

QY 94 TGAGCCACCGCGCGCGCGCGCGCGACACCTAGTTTAAAGGCCCTCTGTGTCGCCG 153
Db 1680 TGAGCCACCGCGCGCGCGCGCGCGACCTAGTTTAAAGGCCCTCTGTGTCGCCG 153

QY 154 TGCCTCGCGCTCCCGAGCTGCCAGTC 178
Db 1620 TGAGCCAGGCGTGAGACCACTGCC 1596

RESULT 14

AAK85393/c

ID AAK85393 standard; DNA; 444 BP.

AC AAK85393;

DT 07-NOV-2001 (first entry)

DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:40205.

KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
cytostatic; gene therapy; vaccine; metastasis; ds.

OS Homo sapiens.

XX WO200157192-A2.

XX 09-AUG-2001.

XX 17-JAN-2001; 2001WO-US001354.

XX 31-JAN-2000; 2000US-0179065P.

XX 04-FEB-2000; 2000US-0180628P.

XX 24-FEB-2000; 2000US-0184664P.

XX 02-MAR-2000; 2000US-0186350P.

XX 16-MAR-2000; 2000US-0189874P.

XX 17-MAR-2000; 2000US-0190076P.

XX 18-APR-2000; 2000US-0198123P.

XX 19-MAY-2000; 2000US-0205151P.

XX 07-JUN-2000; 2000US-0209467P.

XX 28-JUN-2000; 2000US-0214886P.

XX 30-JUN-2000; 2000US-0215135P.

XX 07-JUL-2000; 2000US-0216647P.

XX 07-JUL-2000; 2000US-0216880P.

XX 11-JUL-2000; 2000US-0217487P.

XX 11-JUL-2000; 2000US-0217496P.

XX 14-JUL-2000; 2000US-0218290P.

XX 26-JUL-2000; 2000US-0220963P.

XX 26-JUL-2000; 2000US-0220964P.

XX 14-AUG-2000; 2000US-0224518P.

XX 14-AUG-2000; 2000US-0224519P.

XX 14-AUG-2000; 2000US-0225213P.

XX 14-AUG-2000; 2000US-0225214P.

PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225477P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 19-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234597P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 01-NOV-2000; 2000US-0241826P.
PR 06-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.

CC medicines for regulation of their expression and activity, or as targets
CC of gene therapy. The genes are involved in tissue and/or cell
CC regeneration. Membrane proteins, signal transduction-related proteins,
CC transcription-related proteins, disease-related proteins and genes
CC encoding them can be used as indicators for diseases (e.g. osteoporosis,
CC neurological diseases, cancer, tumours. The cDNA may be used to regulate
CC the activity or expression of the encoded protein to treat diseases. The
CC sequence presented is a cDNA of the invention. Note: Some of the sequence
CC data for this patent is not represented in the printed specification, but
CC is based on sequence information supplied by the European Patent Office.

XX
SQ Sequence 2390 BP; 663 A; 460 C; 500 G; 767 T; 0 U; 0 Other;
Query Match 8.3%; Score 82.2; DB 9; Length 2390;
Best Local Similarity 87.4%; Pred. No. 3.1e-10;
Matches 90; Conservative 0; Mismatches 13; Indels 0; Gaps 0;
Qy 34 CGATCTCTGACCTCGTGATCCGCCCGCCTCAGCTTCCCAAAGTGTGGATTACAGGCA 93
Db 2200 CGATCTCTGACCTCGTGATCCGCCCGCCTCAGCTTCCCAAAGTGTGGATTACAGGCG 2141
Qy 94 TGAGCCACCGCGCCCGCCGACACCTAGTTTAAAGGGCC 136
Db 2140 TGAGCCACCGCGCCCGCGACTAAAGGAATTTTAAAGGGC 2098

Search completed: July 10, 2004, 09:54:39
Job time : 502 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 10, 2004, 09:37:04 ; Search time 100 Seconds
(without alignments)
5494.014 Million cell updates/sec

Title: US-09-972-032-1

Perfect score: 990
Sequence: 1 ggaattgtctcgaggccaa.....aaatgagcgccgaagt 990

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents NA:*

1: /cgn2_6/ptodata/2/ina/5A_COMB.seq.*

2: /cgn2_6/ptodata/2/ina/5B_COMB.seq.*

3: /cgn2_6/ptodata/2/ina/6A_COMB.seq.*

4: /cgn2_6/ptodata/2/ina/6B_COMB.seq.*

5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq.*

6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	83.6	8.4	506	4	US-09-621-976-15373
2	80.4	8.1	51719	4	US-09-918-686-2
3	80.4	8.1	92139	4	US-09-918-686-1
4	78.8	8.0	511	4	US-09-621-976-9083
5	78.4	7.9	480	4	US-09-621-976-8284
6	78.4	7.9	21721	4	US-09-269-939A-41
7	78.4	7.9	22976	4	US-09-269-939A-19
8	78.4	7.9	23187	4	US-09-499-522-1
9	77.8	7.9	397	4	US-09-621-976-12615
10	77.6	7.8	2280	4	US-09-702-705-321
11	77.6	7.8	2280	4	US-09-736-457-321
12	77.6	7.8	2280	4	US-09-614-124B-321
13	77.6	7.8	2280	4	US-09-671-325-321
14	77.6	7.8	2280	4	US-09-589-184-321
15	77.6	7.8	319608	4	US-09-539-333D-1
16	77.6	7.8	319608	4	US-09-679-409-1
17	77.2	7.8	162450	4	US-09-345-882-1
18	77	7.8	320	1	US-08-629-939-5
19	77	7.8	320	1	US-08-759-873-5
20	76.8	7.8	1000	4	US-09-671-317-469
21	76.8	7.8	1000	4	US-09-671-317-484
22	76.8	7.8	49312	4	US-09-671-317-485
23	76.8	7.8	786431	4	US-09-751-389-3
24	76.4	7.7	40000	4	US-09-780-049-18
25	76.4	7.7	63588	4	US-09-873-404-3
26	76.2	7.7	578	4	US-09-495-050A-99
27	76.2	7.7	87350	3	US-08-781-891-79

Sequence 79, Appl
Sequence 3, Appli
Sequence 11, Appl
Sequence 37, Appl
Sequence 44, Appl
Sequence 3, Appli
Sequence 12757, A
Sequence 223, App
Sequence 20, Appl
Sequence 21, Appl
Sequence 22, Appl
Sequence 101, App
Sequence 102, App
Sequence 3, Appli
Sequence 652, App
Sequence 12, Appl
Sequence 3, Appli
Sequence 1, Appli

7.7 87350 4 US-09-618-166-79
7.7 87543 4 US-09-791-211-3
7.7 392000 4 US-10-027-983-11
7.7 193303 4 US-09-497-855A-37
7.7 193303 4 US-09-497-855A-44
7.7 193303 4 US-09-077-354B-3
7.7 10380 3 US-09-621-976-12757
7.6 480 4 US-09-671-317-223
7.6 1001 4 US-08-724-394A-20
7.6 246240 2 US-08-724-394A-21
7.6 246240 2 US-08-724-394A-22
7.6 40328 3 US-08-742-185-102
7.6 43795 3 US-08-742-185-101
7.6 116592 4 US-09-818-512-3
7.6 489 4 US-09-621-976-652
7.6 18073 3 US-09-078-294-12
7.6 13208 4 US-09-852-067-3
7.6 81001 4 US-09-750-580-1

ALIGNMENTS

RESULT 1

US-09-621-976-15373

; Sequence 15373, Application US/09621976

; Patent No. 6639063

; GENERAL INFORMATION:

; APPLICANT: Dumas Milne Edwards, J.B.

; APPLICANT: Jobert, S.

; APPLICANT: Giordano, J.Y.

; TITLE OF INVENTION: ESTs and Encoded Human Proteins.

; FILE REFERENCE: GENSET.054PR2

; CURRENT APPLICATION NUMBER: US/09/621,976

; CURRENT FILING DATE: 2000-07-21

; NUMBER OF SEQ ID NOS: 19335

; SOFTWARE: Patent.pm

; SEQ ID NO 15373

; LENGTH: 506

; TYPE: DNA

; ORGANISM: Homo sapiens

; US-09-621-976-15373

Query Match 8.4%; Score 83.6; DB 4; Length 506;
Best Local Similarity 90.8%; Pred. No. 5.8e-12;
Matches 89; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 34 CGATCTCTGACCTCGTGTATCGCCGCTCAGCTTCCAAAGTCTGGGATTACAGCA 93

Db 41 CGATCTCTTACCTCGTGTATCGCCGCTCGGCTCCCAAAGTCTGGGATTACAGCA 100

QY 94 TGAGCCACGGCCCGCCCGCCGACACCTAGTTTAAA 131

Db 101 TGAGCCACGGCCCGCCCGCCCAATCATCTGTTTTAAA 138

RESULT 2

US-09-918-686-2/c

; Sequence 2, Application US/09918686

; Patent No. 6475739

; GENERAL INFORMATION:

; APPLICANT: Brunkow, Mary

; APPLICANT: Prohl, Sean

; APPLICANT: Paepel, Bryan

; APPLICANT: Staehling-Hampton, Karen

; TITLE OF INVENTION: METHODS FOR IDENTIFYING

; FILE REFERENCE: GENOMIC DELETIONS

; CURRENT APPLICATION NUMBER: 240083.515

; CURRENT FILING DATE: 2001-07-30

; NUMBER OF SEQ ID NOS: 105

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 2

```
; LENGTH: 51719
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1246, 2572, 2604
; OTHER INFORMATION: n = A,T,C or G
US-09-918-686-2

Query Match      8.1%; Score 80.4; DB 4; Length 51719;
Best Local Similarity 88.8%; Pred. No. 1.9e-10;
Matches 87; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 34 CGATCTCTGACCTCGTGTATCCGCGCCCTCAGCTTCCCAAAGTGTGGGATTACAGGCA 93
Db 1119 CGATCTCTGACCTCGTGTATCCGCGCCCTCAGCTTCCCAAAGTGTGGGATTACAGGCA 1060

QY 94 TGAGCCACCGCGCGCGCCCGCCGACACCTAGTTTAA 131
Db 1059 CGAGCCACCGCGCGCGCCCGCCGACATATCTATTTTAA 1022

RESULT 3
US-09-918-686-1/c
; Sequence 1, Application US/09918686
; Patent No. 6475739
; GENERAL INFORMATION:
; APPLICANT: Brunkow, Mary
; APPLICANT: Proll, Sean
; APPLICANT: Paepel, Bryan
; APPLICANT: Staehling-Hampton, Karen
; TITLE OF INVENTION: METHODS FOR IDENTIFYING
; FILE REFERENCE: GENOMIC DELETIONS
; FILE REFERENCE: 240083.515
; CURRENT APPLICATION NUMBER: US/09/918,686
; CURRENT FILING DATE: 2001-07-30
; NUMBER OF SEQ ID NOS: 105
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 92139
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 7043, 8369, 8401
; OTHER INFORMATION: n = A,T,C or G
US-09-918-686-1

Query Match      8.1%; Score 80.4; DB 4; Length 92139;
Best Local Similarity 88.8%; Pred. No. 2.4e-10;
Matches 87; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 34 CGATCTCTGACCTCGTGTATCCGCGCCCTCAGCTTCCCAAAGTGTGGGATTACAGGCA 93
Db 6916 CGATCTCTGACCTCGTGTATCCGCGCCCTCAGCTTCCCAAAGTGTGGGATTACAGGCA 6857

QY 94 TGAGCCACCGCGCGCGCCCGCCGACACCTAGTTTAA 131
Db 6856 CGAGCCACCGCGCGCGCCCGCCGACATATCTATTTTAA 6819

RESULT 4
US-09-918-976-9083/c
; Sequence 9083, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
```

```
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 9083
; LENGTH: 511
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-621-976-9083

Query Match      8.0%; Score 78.8; DB 4; Length 511;
Best Local Similarity 84.0%; Pred. No. 9.4e-11;
Matches 89; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

QY 34 CGATCTCTGACCTCGTGTATCCGCGCCCTCAGCTTCCCAAAGTGTGGGATTACAGGCA 93
Db 439 CGATCTCTGACCTCGTGTATCCGCGCCCTCAGCTTCCCAAAGTGTGGGATTACAGGCG 380

QY 94 TGAGCCACCGCGCGCGCCCGCCGACACCTAGTTTAAAGGGCCCT 139
Db 379 TGAGCCACCGCGCGCGCCCGCCGACCTTCTTAAAGGGATGCT 334

RESULT 5
US-09-621-976-8284/c
; Sequence 8284, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 8284
; LENGTH: 480
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-621-976-8284

Query Match      7.9%; Score 78.4; DB 4; Length 480;
Best Local Similarity 88.5%; Pred. No. 1.2e-10;
Matches 85; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 34 CGATCTCTGACCTCGTGTATCCGCGCCCTCAGCTTCCCAAAGTGTGGGATTACAGGCA 93
Db 472 CGATCTCTGACCTCGTGTATCCGCGCCCTCAGCTTCCCAAAGTGTGGGATTACAGGCG 413

QY 94 TGAGCCACCGCGCGCGCCCGCCGACACCTAGTTTAA 129
Db 412 TGAGCCACCGCGCGCGCCCGCCGAGATATATTATA 377

RESULT 6
US-09-269-939A-41
; Sequence 41, Application US/09269939A
; Patent No. 6635431
; GENERAL INFORMATION:
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Bihain, Bernard
; APPLICANT: Yen-Potin, Frances
; TITLE OF INVENTION: LSR Receptor, Its Activity, Its Cloning and Its Application To
; TITLE OF INVENTION: The Diagnosis Prevention And/or Treatment of Obesity and
; TITLE OF INVENTION: Related Risks or Complications
; FILE REFERENCE: GENSET.035APC
; CURRENT APPLICATION NUMBER: US/09/269,939A
; CURRENT FILING DATE: 1999-05-28
; PRIOR APPLICATION NUMBER: FR 97/10088
; PRIOR FILING DATE: 1997-08-06
; PRIOR APPLICATION NUMBER: FR 98/05032
; PRIOR FILING DATE: 1998-04-22
; NUMBER OF SEQ ID NOS: 41
```

```
; SOFTWARE: Patent.pm
; SEQ ID NO 41
; LENGTH: 21721
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: exon
; LOCATION: 1898..2253
; NAME/KEY: exon
; LOCATION: 3438..3782
; NAME/KEY: exon
; LOCATION: 12064..12183
; NAME/KEY: exon
; LOCATION: 15049..15105
; NAME/KEY: exon
; LOCATION: 15670..15816
; NAME/KEY: exon
; LOCATION: 19486..19659
; NAME/KEY: exon
; LOCATION: 19806..19865
; NAME/KEY: exon
; LOCATION: 19963..20094
; NAME/KEY: exon
; LOCATION: 20236..20864
; NAME/KEY: exon
; LOCATION: 20954..21094
; NAME/KEY: Misc_Feature
; LOCATION: 715
; OTHER INFORMATION: diverging nucleotide, G in ref genbank:AC002128
; NAME/KEY: Misc_Feature
; LOCATION: 1229
; OTHER INFORMATION: diverging insertion, G in ref genbank:AC002128
; NAME/KEY: Misc_Feature
; LOCATION: 3676
; OTHER INFORMATION: diverging nucleotide, T in ref genbank:AC002128
; NAME/KEY: Misc_Feature
; LOCATION: 5039
; OTHER INFORMATION: diverging deletion, G in ref genbank:AD000684
; NAME/KEY: Misc_Feature
; LOCATION: 5118
; OTHER INFORMATION: diverging nucleotide, C in ref genbank:AD000684
; NAME/KEY: Misc_Feature
; LOCATION: 7337
; OTHER INFORMATION: diverging deletion, C in ref genbank:AD000684
; NAME/KEY: Misc_Feature
; LOCATION: 8294
; OTHER INFORMATION: diverging nucleotide, G in ref genbank:AD000684
; NAME/KEY: Misc_Feature
; LOCATION: 8604
; OTHER INFORMATION: diverging nucleotide, C in ref genbank:AD000684
; NAME/KEY: Misc_Feature
; LOCATION: 8928
; OTHER INFORMATION: diverging nucleotide, A in ref genbank:AD000684
; NAME/KEY: Misc_Feature
; LOCATION: 9021
; OTHER INFORMATION: diverging nucleotide, G in ref genbank:AD000684
; NAME/KEY: Misc_Feature
; LOCATION: 9851
; OTHER INFORMATION: diverging insertion, GAATGAAA in ref genbank:AD000684
; NAME/KEY: Misc_Feature
; LOCATION: 9878
; OTHER INFORMATION: diverging nucleotide, C in ref genbank:AD000684
; NAME/KEY: Misc_Feature
; LOCATION: 11478
; OTHER INFORMATION: diverging nucleotide, T in ref genbank:AD000684
; NAME/KEY: Misc_Feature
; LOCATION: 11577
; OTHER INFORMATION: diverging deletion, C in ref genbank:AD000684
; NAME/KEY: Misc_Feature
; LOCATION: 11779
; OTHER INFORMATION: diverging nucleotide, T in ref genbank:AD000684
; NAME/KEY: Misc_Feature
; LOCATION: 13411
```

```
; OTHER INFORMATION: diverging deletion, T in ref genbank:AD000684
; NAME/KEY: Misc_Feature
; LOCATION: 13538
; OTHER INFORMATION: diverging nucleotide, G in ref genbank:AD000684
; NAME/KEY: Misc_Feature
; LOCATION: 13896
; OTHER INFORMATION: diverging nucleotide, G in ref genbank:AD000684
; NAME/KEY: Misc_Feature
; LOCATION: 14912
; OTHER INFORMATION: diverging nucleotide, A in ref genbank:AD000684
; NAME/KEY: Misc_Feature
; LOCATION: 16732
; OTHER INFORMATION: diverging nucleotide, C in ref genbank:AD000684
; NAME/KEY: Misc_Feature
; LOCATION: 17169
; OTHER INFORMATION: diverging nucleotide, G in ref genbank:AD000684
; NAME/KEY: Misc_Feature
; LOCATION: 18946
; OTHER INFORMATION: diverging deletion, G in ref genbank:AD000684
; NAME/KEY: Misc_Feature
; LOCATION: 19474
; OTHER INFORMATION: diverging nucleotide, G in ref genbank:AD000684
; NAME/KEY: Misc_Feature
; LOCATION: 20500
; OTHER INFORMATION: diverging deletion, G in ref genbank:AD000684
; NAME/KEY: Misc_Feature
; LOCATION: 20501
; OTHER INFORMATION: diverging deletion, G in ref genbank:AD000684
; NAME/KEY: Misc_Feature
; LOCATION: 20502
; OTHER INFORMATION: diverging deletion, A in ref genbank:AD000684
; NAME/KEY: Misc_Feature
; LOCATION: 21270
; OTHER INFORMATION: diverging nucleotide, G in ref genbank:AD000684
; NAME/KEY: Misc_Feature
; LOCATION: 21356
; OTHER INFORMATION: diverging insertion, T in ref genbank:AD000684
; NAME/KEY: Misc_Feature
; LOCATION: 21476
; OTHER INFORMATION: diverging nucleotide, A in ref genbank:AD000684
; NAME/KEY: Misc_Feature
; LOCATION: 21588
; OTHER INFORMATION: diverging insertion, C in ref genbank:AD000684
; NAME/KEY: Misc_Feature
; LOCATION: 21601
; OTHER INFORMATION: diverging deletion, T in ref genbank:AD000684
; NAME/KEY: Misc_Feature
; LOCATION: 21635
; OTHER INFORMATION: diverging insertion, G in ref genbank:AD000684
; NAME/KEY: Misc_Feature
; LOCATION: 19963..19965
; OTHER INFORMATION: Potential variant splicing site AAG
; NAME/KEY: Misc_Feature
; LOCATION: 1..21721
; OTHER INFORMATION: n= a,g,c or t
; US-09-269-939A-41
; Query Match 7.9%; Score 78.4; DB 4; Length 21721;
; Best Local Similarity 84.6%; Pred. No. 4.6e-10;
; Matches 88; Conservative 0; Mismatches 16; Indels 0; Gaps 0;
QY 34 CGATCTCTGACCTCGTGATCGCCGCTCAGCTTCCCAAGTCTCGGATTACAGCA 93
Db 5178 CGATCTCTGACCTCGTGATCGCCGCTCAGCTTCCCAAGTCTCGGATTACAGCG 5237
QY 94 TGAGCCACCGCGCCGCCGACACCTAGTTTAAAGGCC 137
Db 5238 TGAGCCACCGCGCCGCCGCTTGTCTTAATGTATGCTGCC 5281
; RESULT 7
; US-09-269-939A-19
; Sequence 19, Application US/09269939A
```

```
; Patent No. 6635431
; GENERAL INFORMATION:
; APPLICANT: Bihaïn, Bernard
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Yen-Potin, Frances
; TITLE OF INVENTION: LSR Receptor, Its Activity, Its Cloning and Its Application To
; TITLE OF INVENTION: The Diagnosis Prevention And/or Treatment of Obesity and
; TITLE OF INVENTION: Related Risks or Complications
; FILE REFERENCE: GENSET.035APC
; CURRENT APPLICATION NUMBER: US/09/269,939A
; CURRENT FILING DATE: 1999-05-28
; PRIOR APPLICATION NUMBER: FR 97/10088
; PRIOR FILING DATE: 1997-08-06
; PRIOR APPLICATION NUMBER: FR 98/05032
; PRIOR FILING DATE: 1998-04-22
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: Patent.pm
; SEQ ID NO 19
; LENGTH: 22976
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURES:
; NAME/KEY: exon
; LOCATION: 1898..2253
; NAME/KEY: exon
; LOCATION: 3437..3781
; NAME/KEY: exon
; LOCATION: 12065..12184
; NAME/KEY: exon
; LOCATION: 15045..15101
; NAME/KEY: exon
; LOCATION: 15666..15812
; NAME/KEY: exon
; LOCATION: 19479..19652
; NAME/KEY: exon
; LOCATION: 19799..19858
; NAME/KEY: exon
; LOCATION: 19956..20087
; NAME/KEY: exon
; LOCATION: 20229..20854
; NAME/KEY: exon
; LOCATION: 20944..21094
; NAME/KEY: Misc Feature
; LOCATION: 19956..19958
; OTHER INFORMATION: Potential variant splicing site AAG
; US-09-939A-19

Query Match 7.9% Score 78.4; DB 4; Length 22976;
Best Local Similarity 84.6%; Pred. No. 4.6e-10;
Matches 88; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

Qy 34 CGATCTCTGACCTCGTGATCGCGCGCTCAGCTTCCCAAGTCTGGGATTACAGGCA 93
Db 5176 CGATCTCTGACCTCGTGATCGCGCGCTCGCGCTCCCAAGTCTGGGATTACAGGCG 5235

Qy 94 TGAGCCACCGCGCGCGCGCGCCGACACCTAGTTTAAAGGGCCC 137
Db 5236 TGAGCCACCGCGCGCGCGCGCTTGTCTTAACTGTAATGCTGCC 5279

RESULT 8
US-09-499-522-1
; Sequence 1, Application US/09499522
; Patent No. 6479238
; GENERAL INFORMATION:
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Bihaïn, Bernard
; TITLE OF INVENTION: POLYMORPHIC MARKERS OF THE LSR GENE
; FILE REFERENCE: GENSET.053AUS
; CURRENT APPLICATION NUMBER: US/09/499,522
; CURRENT FILING DATE: 2000-02-10
; EARLIER APPLICATION NUMBER: US 60/119,592
```

```
; EARLIER FILING DATE: 1999-02-10
; EARLIER APPLICATION NUMBER: US 60/144,784
; EARLIER FILING DATE: 1999-07-20
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Patent.pm
; SEQ ID NO 1
; LENGTH: 23187
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURES:
; NAME/KEY: exon
; LOCATION: 2001..2356
; OTHER INFORMATION: exon1
; FEATURE:
; NAME/KEY: exon
; LOCATION: 3540..3884
; OTHER INFORMATION: exon2
; FEATURE:
; NAME/KEY: exon
; LOCATION: 12163..12282
; OTHER INFORMATION: exon3
; FEATURE:
; NAME/KEY: exon
; LOCATION: 15144..15200
; OTHER INFORMATION: exon4
; FEATURE:
; NAME/KEY: exon
; LOCATION: 15765..15911
; OTHER INFORMATION: exon5
; FEATURE:
; NAME/KEY: exon
; LOCATION: 19579..19752
; OTHER INFORMATION: exon6
; FEATURE:
; NAME/KEY: exon
; LOCATION: 19899..19958
; OTHER INFORMATION: exon7
; FEATURE:
; NAME/KEY: exon
; LOCATION: 20056..20187
; OTHER INFORMATION: exon8
; FEATURE:
; NAME/KEY: exon
; LOCATION: 20329..20957
; OTHER INFORMATION: exon9
; FEATURE:
; NAME/KEY: exon
; LOCATION: 21047..21187
; OTHER INFORMATION: exon10
; FEATURE:
; NAME/KEY: polyA signal
; LOCATION: 21168..21173
; OTHER INFORMATION: AATAAA
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 1..2600
; OTHER INFORMATION: potential 5' regulatory region
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 22324..23187
; OTHER INFORMATION: homology with USF2 gene in ref: embl Y07661
; FEATURE:
; NAME/KEY: primer bind
; LOCATION: 523..544
; OTHER INFORMATION: upstream amplification primer 17-2
; FEATURE:
; NAME/KEY: primer bind
; LOCATION: 1047..1068
; OTHER INFORMATION: downstream amplification primer 17-2, complement
; FEATURE:
; NAME/KEY: primer bind
; LOCATION: 946..963
; OTHER INFORMATION: upstream amplification primer 99-4576
```


FEATURE:
NAME/KEY: primer_bind
LOCATION: 1385..1402
OTHER INFORMATION: downstream amplification primer 99-4576 , complement
FEATURE:
NAME/KEY: primer_bind
LOCATION: 1096..1115
OTHER INFORMATION: upstream amplification primer 9-19
FEATURE:
NAME/KEY: primer_bind
LOCATION: 1616..1635
OTHER INFORMATION: downstream amplification primer 9-19 , complement
FEATURE:
NAME/KEY: primer_bind
LOCATION: 1602..1621
OTHER INFORMATION: upstream amplification primer 9-20
FEATURE:
NAME/KEY: primer_bind
LOCATION: 2074..2093
OTHER INFORMATION: downstream amplification primer 9-20 , complement
FEATURE:
NAME/KEY: primer_bind
LOCATION: 2036..2053
OTHER INFORMATION: upstream amplification primer 99-4557
FEATURE:
NAME/KEY: primer_bind
LOCATION: 2563..2580
OTHER INFORMATION: downstream amplification primer 99-4557 , complement
FEATURE:
NAME/KEY: primer_bind
LOCATION: 2084..2102
OTHER INFORMATION: upstream amplification primer 9-1
FEATURE:
NAME/KEY: primer_bind
LOCATION: 2483..2500
OTHER INFORMATION: downstream amplification primer 9-1 , complement
FEATURE:
NAME/KEY: primer_bind
LOCATION: 2470..2489
OTHER INFORMATION: upstream amplification primer 9-21 , complement
FEATURE:
NAME/KEY: primer_bind
LOCATION: 2062..2081
OTHER INFORMATION: downstream amplification primer 9-21
FEATURE:
NAME/KEY: primer_bind
LOCATION: 3455..3474
OTHER INFORMATION: upstream amplification primer 9-3
FEATURE:
NAME/KEY: primer_bind
LOCATION: 3882..3901
OTHER INFORMATION: downstream amplification primer 9-3 , complement
FEATURE:
NAME/KEY: primer_bind
LOCATION: 3775..3792
OTHER INFORMATION: upstream amplification primer 99-4558
FEATURE:
NAME/KEY: primer_bind
LOCATION: 4335..4356
OTHER INFORMATION: downstream amplification primer 99-4558 , complement
FEATURE:
NAME/KEY: primer_bind
LOCATION: 4902..4920
OTHER INFORMATION: upstream amplification primer 99-14419 , complement
FEATURE:
NAME/KEY: primer_bind
LOCATION: 4444..4463
OTHER INFORMATION: downstream amplification primer 99-14419
FEATURE:
NAME/KEY: primer_bind
LOCATION: 6638..6655
OTHER INFORMATION: upstream amplification primer 99-4577
FEATURE:

NAME/KEY: primer_bind
LOCATION: 7072..7089
OTHER INFORMATION: downstream amplification primer 99-4577 , complement
FEATURE:
NAME/KEY: primer_bind
LOCATION: 7995..8012
OTHER INFORMATION: upstream amplification primer 99-4559
FEATURE:
NAME/KEY: primer_bind
LOCATION: 8576..8593
OTHER INFORMATION: downstream amplification primer 99-4559 , complement
FEATURE:
NAME/KEY: primer_bind
LOCATION: 9622..9639
OTHER INFORMATION: upstream amplification primer 99-3148
FEATURE:
NAME/KEY: primer_bind
LOCATION: 10023..10040
OTHER INFORMATION: downstream amplification primer 99-3148 , complement
FEATURE:
NAME/KEY: primer_bind
LOCATION: 9964..9981
OTHER INFORMATION: upstream amplification primer 99-4560
FEATURE:
NAME/KEY: primer_bind
LOCATION: 10546..10563
OTHER INFORMATION: downstream amplification primer 99-4560 , complement
FEATURE:
NAME/KEY: primer_bind
LOCATION: 10996..11015
OTHER INFORMATION: upstream amplification primer 99-14411 , complement
FEATURE:
NAME/KEY: primer_bind
LOCATION: 10492..10512
OTHER INFORMATION: downstream amplification primer 99-14411
FEATURE:
NAME/KEY: primer_bind
LOCATION: 11972..11990
OTHER INFORMATION: upstream amplification primer 99-4561
FEATURE:
NAME/KEY: primer_bind
LOCATION: 12481..12501
OTHER INFORMATION: downstream amplification primer 99-4561 , complement
Query Match 7.9%; Score 78.4; DB 4; Length 23187;
Best Local Similarity 84.6%; Pred. No. 4.7e-10;
Matches 88; Conservative 0; Mismatches 16; Indels 0; Gaps 0;
QY 34 CGATCTCTGACCTCGTGATCCGCCGCTCAGCTTCCCAAGTCTGGGATTACAGCA 93
|||
Db 5280 CGATCTCTGACCTCGTGATCCGCCGCTCAGCTTCCCAAGTCTGGGATTACAGCG 5339
QY 94 TGAGCCACCGCCGCCGCCGACACCTAGTTTAAAGGGCCC 137
|||
Db 5340 TGAGCCACCGCCGCCGCCGCTTGTCTTAATGTAATGCTGCC 5383
RESULT 9
US-09-621-976-12615/c
; Sequence 12615, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET 054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 12615
; LENGTH: 397

```
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-621-976-12615

Query Match      7.9%; Score 77.8; DB 4; Length 397;
Best Local Similarity 87.6%; Pred. No. 1.5e-10;
Matches 85; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 34 CGATCTCTGACCTCGTGATCCGCCCGCTCAGCTTCCCAAAGTCTGGGATTACAGGCA 93
    |||||
Db 194 CGATCTCTGACCTCGTGATCCGCCCGCTCAGCTTCCCAAAGTCTGGGATTACAGGCG 135

QY 94 TGAGCCACCGCGCCGCCCGCCGACACTAGTTTAA 130
    |||||
Db 134 TGAGCCACCGCGCCGCCCGCCGACACTATTTCTTAA 98

RESULT 10
US-09-702-705-321
; Sequence 321, Application US/09702705
; Patent No. 6504010
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fang, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C14
; CURRENT APPLICATION NUMBER: US/09/702,705
; CURRENT FILING DATE: 2000-10-30
; NUMBER OF SEQ ID NOS: 1833
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 321
; LENGTH: 2280
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-702-705-321

Query Match      7.8%; Score 77.6; DB 4; Length 2280;
Best Local Similarity 82.4%; Pred. No. 3.2e-10;
Matches 89; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

QY 34 CGATCTCTGACCTCGTGATCCGCCCGCTCAGCTTCCCAAAGTCTGGGATTACAGGCA 93
    |||||
Db 1081 CGATCTCTGACCTCGTGATCCGCCCGCTCAGCTTCCCAAAGTCTGGGATTACAGGCG 1140

QY 94 TGAGCCACCGCGCCGCCCGCCGACACTAGTTTAAAGGCCCCCTGC 141
    |||||
Db 1141 TGAGCCACCGCGCCGCCCGCTGATTCTCTTAAATTTGAAGAGGTGCTGC 1188

RESULT 11
US-09-736-457-321
; Sequence 321, Application US/09736457
; Patent No. 6509448
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fang, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; APPLICANT: Wang, Aijun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
```

```
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C15
; CURRENT APPLICATION NUMBER: US/09/736,457
; CURRENT FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 1864
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 321
; LENGTH: 2280
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-736-457-321

Query Match      7.8%; Score 77.6; DB 4; Length 2280;
Best Local Similarity 82.4%; Pred. No. 3.2e-10;
Matches 89; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

QY 34 CGATCTCTGACCTCGTGATCCGCCCGCTCAGCTTCCCAAAGTCTGGGATTACAGGCA 93
    |||||
Db 1081 CGATCTCTGACCTCGTGATCCGCCCGCTCAGCTTCCCAAAGTCTGGGATTACAGGCG 1140

QY 94 TGAGCCACCGCGCCGCCCGCCGACACTAGTTTAAAGGCCCCCTGC 141
    |||||
Db 1141 TGAGCCACCGCGCCGCCCGCTGATTCTCTTAAATTTGAAGAGGTGCTGC 1188

RESULT 12
US-09-614-124B-321
; Sequence 321, Application US/09614124B
; Patent No. 6630574
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fang, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C9
; CURRENT APPLICATION NUMBER: US/09/614,124B
; CURRENT FILING DATE: 2001-07-11
; NUMBER OF SEQ ID NOS: 1668
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 321
; LENGTH: 2280
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-614-124B-321

Query Match      7.8%; Score 77.6; DB 4; Length 2280;
Best Local Similarity 82.4%; Pred. No. 3.2e-10;
Matches 89; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

QY 34 CGATCTCTGACCTCGTGATCCGCCCGCTCAGCTTCCCAAAGTCTGGGATTACAGGCA 93
    |||||
Db 1081 CGATCTCTGACCTCGTGATCCGCCCGCTCAGCTTCCCAAAGTCTGGGATTACAGGCG 1140

QY 94 TGAGCCACCGCGCCGCCCGCCGACACTAGTTTAAAGGCCCCCTGC 141
    |||||
Db 1141 TGAGCCACCGCGCCGCCCGCTGATTCTCTTAAATTTGAAGAGGTGCTGC 1188

RESULT 13
US-09-671-325-321
; Sequence 321, Application US/09671325
; Patent No. 6667154
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fang, Gary
```

```
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C12
; CURRENT APPLICATION NUMBER: US/09/671,325
; CURRENT FILING DATE: 2000-09-26
; NUMBER OF SEQ ID NOS: 1825
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 321
; LENGTH: 2280
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-671-325-321

Query Match      7.8%; Score 77.6; DB 4; Length 2280;
Best Local Similarity 82.4%; Pred. No. 3.2e-10;
Matches 89; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

Qy      34  CGATCTCTGACCTCGTGATCCGCGCCCTCAGCTTCCCAAAGTGCTGGGATTACAGGCA 93
Db      1081 CGATCTCTGACCTCGTGATCCGCGCCCTCAGCTTCCCAAAGTGCTGGGATTACAGGCG 1140

Qy      94  TGAGCCACCGCGCCCGCCCGCCGACACCTAGTTTAAAGGCCCTGTC 141
Db      1141 TGAGCCACCGTGGCGCGCTGATTCTCTTAAATTTGAAGAGGTGCTGC 1188

RESULT 14
US-09-589-184-321
; Sequence 321, Application US/09589184
; Patent No. 6686447
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C8
; CURRENT APPLICATION NUMBER: US/09/589,184
; CURRENT FILING DATE: 2000-06-05
; NUMBER OF SEQ ID NOS: 827
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 321
; LENGTH: 2280
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-589-184-321

Query Match      7.8%; Score 77.6; DB 4; Length 2280;
Best Local Similarity 82.4%; Pred. No. 3.2e-10;
Matches 89; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

Qy      34  CGATCTCTGACCTCGTGATCCGCGCCCTCAGCTTCCCAAAGTGCTGGGATTACAGGCA 93
Db      1081 CGATCTCTGACCTCGTGATCCGCGCCCTCAGCTTCCCAAAGTGCTGGGATTACAGGCG 1140

Qy      94  TGAGCCACCGCGCCCGCCCGCCGACACCTAGTTTAAAGGCCCTGTC 141
Db      1141 TGAGCCACCGTGGCGCGCTGATTCTCTTAAATTTGAAGAGGTGCTGC 1188

RESULT 15
US-09-539-333D-1/c
; Sequence 1, Application US/09539333D
```

```
; Patent No. 6476208
; GENERAL INFORMATION:
; APPLICANT: Cohen, Daniel
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Chumakov, Ilya
; APPLICANT: Boumleheret, Lydie
; APPLICANT: Bihain, Bernard
; APPLICANT: Essioux, Laurent
; TITLE OF INVENTION: SCHIZOPHRENIA ASSOCIATED GENES, PROTEINS AND BIALLELIC MARKERS
; FILE REFERENCE: GENSET.047AUS
; CURRENT APPLICATION NUMBER: US/09/539,333D
; CURRENT FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: US 60/126,903
; PRIOR FILING DATE: 1999-03-30
; PRIOR APPLICATION NUMBER: US 60/131,971
; PRIOR FILING DATE: 1999-04-30
; PRIOR APPLICATION NUMBER: US 60/132,065
; PRIOR FILING DATE: 1999-04-30
; PRIOR APPLICATION NUMBER: US 60/143,928
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: US 60/145,915
; PRIOR FILING DATE: 1999-07-27
; PRIOR APPLICATION NUMBER: US 60/146,453
; PRIOR FILING DATE: 1999-07-29
; PRIOR APPLICATION NUMBER: US 60/146,452
; PRIOR FILING DATE: 1999-07-29
; PRIOR APPLICATION NUMBER: US 60/162,288
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: US 09/416,384
; PRIOR FILING DATE: 1999-10-12
; NUMBER OF SEQ ID NOS: 231
; SOFTWARE: Patent.pm
; SEQ ID NO 1
; LENGTH: 319608
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 31..1107
; OTHER INFORMATION: 5'regulatory region g35018 gene
; NAME/KEY: exon
; LOCATION: 1108..1289
; OTHER INFORMATION: exon A g35018 gene
; FEATURE:
; NAME/KEY: exon
; LOCATION: 14877..14920
; OTHER INFORMATION: exon B g35018 gene
; FEATURE:
; NAME/KEY: exon
; LOCATION: 18778..18862
; OTHER INFORMATION: exon Bbis g35018 gene
; FEATURE:
; NAME/KEY: exon
; LOCATION: 25593..25740
; OTHER INFORMATION: exon C g35018 gene
; FEATURE:
; NAME/KEY: exon
; LOCATION: 29388..29502
; OTHER INFORMATION: exon D g35018 gene
; FEATURE:
; NAME/KEY: exon
; LOCATION: 29967..30282
; OTHER INFORMATION: exon E g35018 gene
; FEATURE:
; NAME/KEY: exon
; LOCATION: 64666..64812
; OTHER INFORMATION: exon F g35018 gene
; FEATURE:
; NAME/KEY: exon
; LOCATION: 65505..65853
; OTHER INFORMATION: exon G g35018 gene
; FEATURE:
```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 10, 2004, 11:04:44 ; Search time 537 Seconds

(Without alignments)
8985.372 Million cell updates/sec

Title: US-09-972-032-1

Perfect score: 990

Sequence: 1 ggaattgtcttcgagccaa.....aaaatgagcgcgcgaagt 990

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3183909 seqs, 2436941669 residues

Total number of hits satisfying chosen parameters: 6367818

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq2:*
- 14: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
- 16: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
- 17: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
- 18: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
- 19: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	990	100.0	990	13	US-09-972-032-1
2	86.6	8.7	631	13	Sequence 1, Appl1
3	86.6	8.7	631	16	Sequence 251882,
4	85	8.6	96597	16	Sequence 251882,
5	84.8	8.6	99291	17	Sequence 112, App
6	84.2	8.5	67783	12	Sequence 744, App
7	83.6	8.4	2105	13	Sequence 238, App
8	83.6	8.4	325348	16	Sequence 330, App
9	83.2	8.4	636	13	Sequence 358, App
10	83.2	8.4	636	16	Sequence 269840,
11	83	8.4	598	13	Sequence 269840,
12	83	8.4	598	16	Sequence 115506,
13	83	8.4	13695	16	Sequence 115506,
14	83	8.4	65387	17	Sequence 905, App
					Sequence 332, App

15	82.8	8.4	790	13	US-10-027-632-125805	Sequence 125805,
16	82.8	8.4	790	13	US-10-027-632-125806	Sequence 125806,
17	82.8	8.4	790	16	US-10-027-632-125805	Sequence 125805,
18	82.8	8.4	790	16	US-10-027-632-125806	Sequence 125806,
19	82.2	8.3	741	13	US-10-027-632-149163	Sequence 149163,
20	82.2	8.3	2114	13	US-10-027-632-149163	Sequence 149163,
21	82.2	8.3	2114	13	US-10-027-632-97278	Sequence 97278, A
22	82.2	8.3	2114	16	US-10-027-632-97278	Sequence 97278, A
23	82.2	8.3	2390	16	US-10-104-047-1390	Sequence 1390, Ap
24	82.2	8.3	12477	9	US-09-764-870-559	Sequence 559, App
25	82.2	8.3	12477	9	US-09-764-853-851	Sequence 851, App
26	82.2	8.3	12477	15	US-10-125-540-559	Sequence 559, App
27	82.2	8.3	12477	15	US-10-103-313-579	Sequence 579, App
28	82.2	8.3	12477	16	US-10-158-057-339	Sequence 339, App
29	82.2	8.3	26332	13	US-10-087-192-1558	Sequence 1558, Ap
30	82	8.3	636	13	US-10-027-632-269839	Sequence 269839,
31	82	8.3	636	16	US-10-027-632-269839	Sequence 269839,
32	82	8.3	2494	16	US-10-108-260A-127	Sequence 127, App
33	82	8.3	2545	13	US-10-027-632-112152	Sequence 112152,
34	82	8.3	2545	13	US-10-027-632-112153	Sequence 112153,
35	82	8.3	2545	16	US-10-027-632-112152	Sequence 112152,
36	82	8.3	2545	16	US-10-027-632-112153	Sequence 112153,
37	82	8.3	2904	16	US-10-104-047-1356	Sequence 1356, Ap
38	82	8.3	16491	10	US-09-764-891-5505	Sequence 5505, Ap
39	81.8	8.3	16821	13	US-10-087-192-622	Sequence 622, App
40	81.6	8.2	387	9	US-09-867-701-10757	Sequence 10757, A
41	81.6	8.2	636	13	US-10-027-632-275829	Sequence 275829,
42	81.6	8.2	636	16	US-10-027-632-275829	Sequence 275829,
43	81.2	8.2	351	9	US-09-867-701-9314	Sequence 9314, Ap
44	81.2	8.2	361	9	US-09-867-701-9326	Sequence 9326, Ap
45	81.2	8.2	389	9	US-09-867-701-10783	Sequence 10783, A

ALIGNMENTS

RESULT 1
US-09-972-032-1
; Sequence 1, Application US/09972032
; Publication No. US20020086361A1
; GENERAL INFORMATION:
; APPLICANT: Case Western Reserve University
; APPLICANT: Montano, Monica
; APPLICANT: Sutton, Amelia
; TITLE OF INVENTION: A Modulator of Antiestrogen Pharmacology
; FILE REFERENCE: 27708/04003
; CURRENT APPLICATION NUMBER: US/09/972,032
; CURRENT FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: US 60/238,190
; PRIOR FILING DATE: 2000-10-05
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 1
; LENGTH: 990
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-972-032-1

Query Match	100.0%	Score 990;	DB 13;	Length 990;
Best Local Similarity	100.0%	Pred. No. 2.4e-269;		
Matches 990;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	GGAAATGTTCTCGAGCCCAAGAAATTCGGCAAGCGGATCTCTGACCTCGTATCCGCGG	60	
Db	1	GGAAATGTTCTCGAGCCCAAGAAATTCGGCAAGCGGATCTCTGACCTCGTATCCGCGG	60	
QY	61	CCTCAGCTTCCCAAAAGTGTGGGATTACAGGCATGAGCACCGCGCCCGCCCGACAC	120	
Db	61	CCTCAGCTTCCCAAAAGTGTGGGATTACAGGCATGAGCACCGCGCCCGCCCGACAC	120	
QY	121	CTAGTTTAAAGGGCCCTGCTGTGTCGCGCTCCAGTCCAGTCCAGTCTG	180	
Db	121	CTAGTTTAAAGGGCCCTGCTGTGTCGCGCTCCAGTCCAGTCCAGTCTG	180	

```
QY 181 CCGGGCTCAGTCCCGCGTTGCCATGTGTGGAGACCCGCTGCGTAAAGCCTGGATGTGG 240
Db 181 CCGGGCTCAGTCCCGCGTTGCCATGTGTGGAGACCCGCTGCGTAAAGCCTGGATGTGG 240
QY 241 CTTGCTGATGACATATGGACCGGGCTCTGAGCTAGGCTAGGGAAGGGCAGAGGGCGG 300
Db 241 CTTGCTGATGACATATGGACCGGGCTCTGAGCTAGGCTAGGGAAGGGCAGAGGGCGG 300
QY 301 AATTGGGCCCCAGGGCCAGAGGCTTCGCCGACCCCCGACTCGCCCTCCGGTGGCCCCGAC 360
Db 301 AATTGGGCCCCAGGGCCAGAGGCTTCGCCGACCCCCGACTCGCCCTCCGGTGGCCCCGAC 360
QY 361 CGCTCCCGGTGGCCCTGGAGTGCAGGTCTTACCGTCCGAGATCGTCCGCAACTGGGCGA 420
Db 361 CGCTCCCGGTGGCCCTGGAGTGCAGGTCTTACCGTCCGAGATCGTCCGCAACTGGGCGA 420
QY 421 GCTGTGATGGGGCGGTGGCTAAAGCCCGTGGTTTGGTTACGATTGGCCAGCGGACTTAAG 480
Db 421 GCTGTGATGGGGCGGTGGCTAAAGCCCGTGGTTTGGTTACGATTGGCCAGCGGACTTAAG 480
QY 481 TGTGTCTCTGAAGAGCATGACATTAGTCTGAGGGTCTCTGGAAGTGTATCCCGCCC 540
Db 481 TGTGTCTCTGAAGAGCATGACATTAGTCTGAGGGTCTCTGGAAGTGTATCCCGCCC 540
QY 541 CACCATCAATGGCGCTTAGTCTAGGAAGCGGGTGTGGGTGGGGCCCTTAGGGCGAGGGG 600
Db 541 CACCATCAATGGCGCTTAGTCTAGGAAGCGGGTGTGGGTGGGGCCCTTAGGGCGAGGGG 600
QY 601 CAGACATACCCGGAAGTGGTTGATTTATACGCAAGGGGCTGGATCGAACCCGCCAAA 660
Db 601 CAGACATACCCGGAAGTGGTTGATTTATACGCAAGGGGCTGGATCGAACCCGCCAAA 660
QY 661 GACACTGAAGGCTGTGTGGCTGAGGAGGCGCGCAATCCAGTGTGTGTGGGCTTTAC 720
Db 661 GACACTGAAGGCTGTGTGGCTGAGGAGGCGCGCAATCCAGTGTGTGTGGGCTTTAC 720
QY 721 AGGAAGAGCTCCACCTTCTTGGAGTGTGCAGATGCGATAGGTGTGTCCACCGATGG 780
Db 721 AGGAAGAGCTCCACCTTCTTGGAGTGTGCAGATGCGATAGGTGTGTGTCCACCGATGG 780
QY 781 GAGCTGGGGCGGGCAGAGTGTGCCCCAGTACAGAGCTGATTTGACCTGGGGCTCTG 840
Db 781 GAGCTGGGGCGGGCAGAGTGTGCCCCAGTACAGAGCTGATTTGACCTGGGGCTCTG 840
QY 841 GACTTCCTGATCTCTGCTGTGATCTCCAGCAAGTCTCTGTCCTCCGTTGGCTGCTTCA 900
Db 841 GACTTCCTGATCTCTGCTGTGATCTCCAGCAAGTCTCTGTCCTCCGTTGGCTGCTTCA 900
QY 901 CCATCTCTCACTTCTCTGCTTCCAGAGTAAATTCAGAGTCTGTGGTGCAAAAAA 960
Db 901 CCATCTCTCACTTCTCTGCTTCCAGAGTAAATTCAGAGTCTGTGGTGCAAAAAA 960
QY 961 AAAAAAATAAATAGCGCGCGCAAGTT 990
Db 961 AAAAAAATAAATAGCGCGCGCAAGTT 990
```

RESULT 2

```
US-10-027-632-251882
; Sequence 251882, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
```

```
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 251882
; LENGTH: 631
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-251882
```

```
Query Match      8.7%; Score 86.6; DB 13; Length 631;
Best Local Similarity 91.1%; Pred. No. 3.3e-14;
Matches 92; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 34 CGATCTCTGACCTGCTGATCCGCGCCGCTCAGCTTCCCAAAGTCTGGGATTACAGGCA 93
Db 409 CGATCTCTGACCTGCTGATCCGCGCCGCTCAGCTTCCCAAAGTCTGGGATTACAGGCT 468
QY 94 TGAGCCACCGCGCCGCGCCGCGCCGACACCTAGTTTAAAGGG 134
Db 469 TGAGCCACCGCGCCGCGCCGCGCCGACCTGACAGCAGTTTAAATGG 509
```

RESULT 3

```
US-10-027-632-251882
; Sequence 251882, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 251882
; LENGTH: 631
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-251882
```

```
Query Match      8.7%; Score 86.6; DB 16; Length 631;
Best Local Similarity 91.1%; Pred. No. 3.3e-14;
Matches 92; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 34 CGATCTCTGACCTGCTGATCCGCGCCGCTCAGCTTCCCAAAGTCTGGGATTACAGGCA 93
Db 409 CGATCTCTGACCTGCTGATCCGCGCCGCTCAGCTTCCCAAAGTCTGGGATTACAGGCT 468
QY 94 TGAGCCACCGCGCCGCGCCGCGCCGACACCTAGTTTAAAGGG 134
```

Db 469 TGAGCCACCGCGCGCCGCACTGACAAGCAGTTTAAATGG 509
|||||

RESULT 4

US-10-085-117-112
; Sequence 112, Application US/10085117
; Publication No. US20030232334A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR CANCER
; FILE REFERENCE: 529452000121
; CURRENT APPLICATION NUMBER: US/10/085,117
; CURRENT FILING DATE: 2002-02-27
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 361
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 112
; LENGTH: 96597
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-085-117-112

Query Match 8.6%; Score 85; DB 16; Length 96597;
Best Local Similarity 77.4%; Pred. No. 4.5e-13;
Matches 103; Conservative 0; Mismatches 30; Indels 0; Gaps 0;
Qy 34 CGATCTCTGACCTCGTATCGCGCCGCTCAGCTTCCCAAAGTGTGGGATTACAGGCA 93
|||
Db 3873 CGATCTCTGACCTCGTATCGCGCCGCTCAGCTTCCCAAAGTGTGGGATTACAGGCG 3932
|||
Qy 94 TGAGCCACCGCGCGCCGCGCCGACACCTAGTTTAAAGGCGCCCTGCTTGTCTCCGC 153
|||
Db 3933 CGAGCCACCGCGCGCCGCGCTCTAGCTTACTTTGAACAATGCCAGAACTCCAGATC 3992
|||
Qy 154 TGC CGCGCTCC 166
|||
Db 3993 CCACACCGCTCC 4005
|||

RESULT 5

US-10-322-281-744
; Sequence 744, Application US/10322281
; Publication No. US20040126762A1
; GENERAL INFORMATION:
; APPLICANT: David W. Morris
; APPLICANT: Marc S. Malandro
; TITLE OF INVENTION: Novel Compositions and Methods in Cancer
; FILE REFERENCE: 529452001000
; CURRENT APPLICATION NUMBER: US/10/322,281
; CURRENT FILING DATE: 2002-12-17
; NUMBER OF SEQ ID NOS: 866
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 744
; LENGTH: 99291
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-322-281-744

Query Match 8.6%; Score 84.8; DB 17; Length 99291;
Best Local Similarity 70.6%; Pred. No. 5.1e-13;
Matches 113; Conservative 0; Mismatches 47; Indels 0; Gaps 0;
Qy 35 GATCTCTGACCTCGTATCGCGCCGCTCAGCTTCCCAAAGTCTGGGATTACAGGCAT 94
|||
Db 20756 GATCTCTGACCTCATGATCGCGCCGCTCAGCTTCCCAAAGTCTGACATTACAGCGGT 20815
|||
Qy 95 GAGCCACCGCGCGCGCCGCGCACACCTAGTTTAAAGGCGCCCTGCTGTGCTGCGCGT 154
|||
Db 20816 GAGCCACCGCGCGCGCGCGCAATCACTATTTTGTAGGGTTTTTTTGTGAGAT 20875
|||

Qy 155 GCCCGCGCTCCAGCTGCCAGTCTGGCGGCTAGTCCC 194
|||
Db 20876 GGACTCTCGCTCTGTTGCCCATGTGGAGTGCAGTGGCAC 20915
|||

RESULT 6

US-09-997-722-238
; Sequence 238, Application US/09997722
; Publication No. US20040072154A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David
; APPLICANT: Engelhard, Eric
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR CANCER
; FILE REFERENCE: A-71171/RMS/DCF
; CURRENT APPLICATION NUMBER: US/09/997,722
; CURRENT FILING DATE: 2001-11-30
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 301
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 238
; LENGTH: 67783
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (34024)..(34043)
; OTHER INFORMATION: "n" at positions 34024 through 34043 can be any base.
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (39350)..(39564)
; OTHER INFORMATION: "n" at positions 39350 through 39564 can be any base.
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (50784)..(50917)
; OTHER INFORMATION: "n" at positions 50784 through 50917 can be any base.
US-09-997-722-238

Query Match 8.5%; Score 84.2; DB 12; Length 67783;
Best Local Similarity 81.0%; Pred. No. 6.8e-13;
Matches 98; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

Qy 34 CGATCTCTGACCTCGTATCGCGCCGCTCAGCTTCCCAAAGTCTGGGATTACAGGCA 93
|||
Db 51742 CGATCTCTGACCTCGTATCGCGCCGCTCAGCTTCCCAAAGTCTGGGATTACAGGCG 51801
|||

Qy 94 TGAGCCACCGCGCGCGCCGCGCACACCTAGTTTAAAGGCGCCCTGCTGTGCTGCGCGC 153
|||
Db 51802 TGAGCCACCGCGCGCGCCGCGCACCTAGTTTAAAGGCGCCCTGCTGTGCTGCGCGC 51861
|||

Qy 154 T 154

Db 51862 T 51862

RESULT 7

US-10-296-115-330/c
; Sequence 330, Application US/10296115
; Publication No. US20040053248A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq Inc
; TITLE OF INVENTION: NO. US20040053248A1el Nucleic Acids and Polypeptides
; FILE REFERENCE: 784PCT
; CURRENT APPLICATION NUMBER: US/10/296,115
; CURRENT FILING DATE: 2002-11-18
; PRIOR APPLICATION NUMBER: US09/488,725
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US09/552,317
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 1478
; SEQ ID NO 330

```
; LENGTH: 2105
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-296-115-330

Query Match      8.4%; Score 83.6; DB 13; Length 2105;
Best Local Similarity 90.8%; Pred. No. 3.4e-13;
Matches 89; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 34 CGATCTCTGACCTCGTGATCCGCGCCGCTCAGCTTCCCAAAGTGTGGGATTACAGGCA 93
    |||||
Db 555 CGATCTCTCTACCTCGTGATCCGCGCCGCTCGGCTCCCAAAGTGTGGGATTACAGGCA 496
    |||||

QY 94 TGAGCCACCGCGCGCGCGCCGCGCCGACACCTAGTTTTTAA 131
    |||||
Db 495 TGAGCCACCGCGCGCGCGCCGCGCCCAATCACTGTTTTTAA 458

RESULT 8
US-10-085-117-358/c
; Sequence 358, Application US/10085117
; Publication No. US20030232334A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR CANCER
; FILE REFERENCE: 529452000121
; CURRENT FILING DATE: 2002-02-27
; PRIOR APPLICATION NUMBER: US/10/085,117
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 361
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 358
; LENGTH: 325348
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: variation
; LOCATION: (1)...(325348)
; OTHER INFORMATION: n = any nucleotide
US-10-085-117-358

Query Match      8.4%; Score 83.6; DB 16; Length 325348;
Best Local Similarity 90.8%; Pred. No. 1.6e-12;
Matches 89; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 34 CGATCTCTGACCTCGTGATCCGCGCCGCTCAGCTTCCCAAAGTGTGGGATTACAGGCA 93
    |||||
Db 156386 CGATCTCTGACCTCGTGATCCGCGCCGCTCAGCTTCCCAAAGTGTGGGATTACAGGCA 156327

QY 94 TGAGCCACCGCGCGCGCGCCGCGCCGACACCTAGTTTTTAA 131
    |||||
Db 156326 TGAGCCACCGCGCGCGCGCCCTTCTTTATTTTAA 156289

RESULT 9
US-10-027-632-269840/c
; Sequence 269840, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 269840
; LENGTH: 636
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-269840

Query Match      8.4%; Score 83.2; DB 16; Length 636;
Best Local Similarity 89.8%; Pred. No. 3.1e-13;
Matches 88; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

QY 34 CGATCTCTGACCTCGTGATCCGCGCCGCTCAGCTTCCCAAAGTGTGGGATTACAGGCA 93
    |||||
Db 350 CGATCTCTCTACCTCGTGATCCGCGCCGCTCGGCTCCCAAAGTGTGGGATTACAGGCA 291
    |||||

QY 94 TGAGCCACCGCGCGCGCGCCGCGCCGACACCTAGTTTTTAA 131
    |||||
Db 290 TGAGCCACCGCGCGCGCCGCGCCCAATCACTGTTTTTAA 253
```

```
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 269840
; LENGTH: 636
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-269840

Query Match      8.4%; Score 83.2; DB 13; Length 636;
Best Local Similarity 89.8%; Pred. No. 3.1e-13;
Matches 88; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

QY 34 CGATCTCTGACCTCGTGATCCGCGCCGCTCAGCTTCCCAAAGTGTGGGATTACAGGCA 93
    |||||
Db 350 CGATCTCTCTACCTCGTGATCCGCGCCGCTCGGCTCCCAAAGTGTGGGATTACAGGCA 291
    |||||

QY 94 TGAGCCACCGCGCGCGCGCCGCGCCGACACCTAGTTTTTAA 131
    |||||
Db 290 TGAGCCACCGCGCGCGCCGCGCCCAATCACTGTTTTTAA 253

RESULT 10
US-10-027-632-269840/c
; Sequence 269840, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 269840
; LENGTH: 636
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-269840

Query Match      8.4%; Score 83.2; DB 16; Length 636;
Best Local Similarity 89.8%; Pred. No. 3.1e-13;
Matches 88; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

QY 34 CGATCTCTGACCTCGTGATCCGCGCCGCTCAGCTTCCCAAAGTGTGGGATTACAGGCA 93
    |||||
Db 350 CGATCTCTCTACCTCGTGATCCGCGCCGCTCGGCTCCCAAAGTGTGGGATTACAGGCA 291
    |||||

QY 94 TGAGCCACCGCGCGCGCGCCGCGCCGACACCTAGTTTTTAA 131
    |||||
Db 290 TGAGCCACCGCGCGCGCCGCGCCCAATCACTGTTTTTAA 253
```



```

; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 115506
; LENGTH: 598
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-115506

Query Match      8.4%; Score 83; DB 16; Length 598;
Best Local Similarity 89.9%; Pred.No. 3.4e-13;
Matches 89; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: Polymorphisms in the Human Genome
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 115506
; LENGTH: 598
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-115506

Query Match      8.4%; Score 83; DB 13; Length 598;
Best Local Similarity 89.9%; Pred.No. 3.4e-13;
Matches 89; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: Polymorphisms in the Human Genome
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 115506
; LENGTH: 598
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-115506

Query Match      8.4%; Score 83; DB 16; Length 598;
Best Local Similarity 89.9%; Pred.No. 3.4e-13;
Matches 89; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

; APPLICANT: SUWA, MAKIKO
; APPLICANT: ASAI, KIYOSHI
; APPLICANT: AKIYAMA, YUTAKA
; APPLICANT: ABURATANI, HIROYUKI
; TITLE OF INVENTION: GUANOSINE TRIPHOSPHATE-BINDING PROTEIN COUPLED RECEPTORS
; FILE REFERENCE: 084335/166
; CURRENT APPLICATION NUMBER: US/10/292,798
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: 10/017,161
; PRIOR FILING DATE: 2001-12-18
; PRIOR APPLICATION NUMBER: JP 2001-246789
; PRIOR FILING DATE: 2001-06-18
; NUMBER OF SEQ ID NOS: 2070
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 905
; LENGTH: 13695
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; LOCATION: source
; FEATURE:
; LOCATION: (1)..(13695)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (201)..(264)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1113)..(1535)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (3116)..(3248)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (3899)..(4046)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (5312)..(5443)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (5789)..(5905)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (6139)..(6184)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (6427)..(6530)
; FEATURE:

```

```

; NAME/KEY: CDS
; LOCATION: (7286) .. (7404)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (9308) .. (9594)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (10681) .. (10789)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (11610) .. (11878)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (13365) .. (13495)
; LOCATION: (13365) .. (13495)
US-10-292-798-905

Query Match      8.4%; Score 83; DB 16; Length 13695;
Best Local Similarity 89.9%; Pred. No. 9e-13;
Matches 89; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 34 CGATCTCTGACCTCGTGATCCGCCCGCCCTCAGCTTCCCAAGTGTGGGATTACAGGCA 93
Db 842 CGATCTCTGACCTCGTGATCCGCCCGCCCTCAGCTTCCCAAGTGTGGGATTACAGGCG 901

QY 94 TGAGCCACCGCGCCCGCCCGCCGACCTAGTTTAAAG 132
Db 902 TGAGCCACCGCGCCCGCCCGCCGACCTAGTTTCTCTTTTATAG 940

RESULT 14
US-10-322-281-332
; Sequence 332, Application US/10322281
; Publication No. US20040126762A1
; GENERAL INFORMATION:
; APPLICANT: David W. Morris
; APPLICANT: Marc S. Malandro
; TITLE OF INVENTION: Novel Compositions and Methods in Cancer
; FILE REFERENCE: 529452001000
; CURRENT APPLICATION NUMBER: US/10/322,281
; CURRENT FILING DATE: 2002-12-17
; NUMBER OF SEQ ID NOS: 866
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 332
; LENGTH: 65387
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-322-281-332

```

```

Query Match      8.4%; Score 83; DB 17; Length 65387;
Best Local Similarity 89.9%; Pred. No. 1.5e-12;
Matches 89; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 34 CGATCTCTGACCTCGTGATCCGCCCGCCCTCAGCTTCCCAAGTGTGGGATTACAGGCA 93
Db 41156 CGATCTCTGACCTCGTGATCCGCCCGCCCTCAGCTTCCCAAGTGTGGGATTACAGGCG 41215

QY 94 TGAGCCACCGCGCCCGCCCGCCGACCTAGTTTAAAG 132
Db 41216 TGAGCCACCGCGCCCGCCCGCCGACCTAGTTTCTCTTTTATAG 41254

```

```

RESULT 15
US-10-027-632-125805
; Sequence 125805, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006

```

```

; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 125805
; LENGTH: 790
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-125805

Query Match      8.4%; Score 82.8; DB 13; Length 790;
Best Local Similarity 88.8%; Pred. No. 4.2e-13;
Matches 87; Conservative 2; Mismatches 9; Indels 0; Gaps 0;

QY 34 CGATCTCTGACCTCGTGATCCGCCCGCCCTCAGCTTCCCAAGTGTGGGATTACAGGCA 93
Db 53 CGATCTCTGACCTCGTGATCCGCCCGCCCTCAGCTTCCCAAGTGTGGGATTACAGGCA 112

QY 94 TGAGCCACCGCGCCCGCCCGCCGACCTAGTTTAAAG 131
Db 113 TGAGCCACCGCGCCCGCCCGCCGACCTAGTTTCTCTTTTAAAG 150

Search completed: July 10, 2004, 13:18:34
Job time : 541 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 10, 2004, 09:31:54 ; Search time 311.0 Seconds
(without alignments)
9505.967 Million cell updates/sec

Title: US-09-972-032-1
Perfect score: 990
Sequence: 1 ggaatggtctcagggccaa.....aaaatgagcgccgaagt 990

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

ZST.*

1: em_estba.*
2: em_esthum.*
3: em_estim.*
4: em_estmu.*
5: em_estov.*
6: em_estpl.*
7: em_estro.*
8: em_htc.*
9: gb_estl.*
10: gb_est2.*
11: gb_htc.*
12: gb_est3.*
13: gb_est4.*
14: gb_est5.*
15: em_estfun.*
16: em_estom.*
17: em_gss_hum.*
18: em_gss_inv.*
19: em_gss_plr.*
20: em_gss_vrt.*
21: em_gss_fun.*
22: em_gss_mam.*
23: em_gss_mus.*
24: em_gss_pro.*
25: em_gss_rod.*
26: em_gss_phg.*
27: em_gss_vrl.*
28: gb_gss1.*
29: gb_gss2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	904.6	91.4	1467	11 BC039500	BC039500 Homo sapi
C 2	793.8	80.2	1201	13 BX406138	BX406138 BX406138
C 3	719.2	72.6	760	12 BM980506	BM980506 UI-CF-EN1
C 4	651.2	65.8	691	13 BU620427	BU620427 UI-H-FL1-

C 5	630.8	63.7	694	13 BU684594	BU684594 UI-CF-EN1
C 6	628.2	63.5	669	12 BM679577	BM679577 UI-E-E00-
C 7	603.2	60.9	632	13 BU617331	BU617331 UI-H-DF0-
C 8	602.6	60.9	633	14 CA313142	CA313142 UI-CF-FN0
C 9	587.8	59.4	623	14 CD742581	CD742581 UI-H-F12-
C 10	576.2	58.2	628	13 BQ447041	BQ447041 UI-H-EU1-
C 11	449.6	45.4	480	10 BF509528	BF509528 UI-H-B14-
C 12	428.4	43.3	462	12 BM687366	BM687366 UI-E-C01-
C 13	409.8	41.4	439	10 BE677246	BE677246 qd2e01.x
C 14	382.8	38.7	411	9 A1184808	A1184808 qd24g11.x
C 15	371	37.5	395	9 AA687318	AA687318 nv61e01.s
C 16	328	33.1	367	10 BF002915	BF002915 7950h12.x
C 17	272.6	27.5	1084	12 BM806108	BM806108 AGENCOURT
C 18	269.6	27.2	300	12 BM672504	BM672504 UI-E-C01-
C 19	211	21.3	282	10 BF364571	BF364571 PMI-NN108
C 20	209	21.1	211	9 A1202243	A1202243 qs65d12.x
C 21	87.4	8.8	698	10 BF663723	BF663723 602145445
C 22	85.8	8.7	456	12 BI792391	BI792391 ic30a05.x
C 23	85.8	8.7	479	28 AQ260685	AQ260685 CITBI-E1-
C 24	85.8	8.7	804	14 CK000918	CK000918 AGENCOURT
C 25	85	8.6	295	10 BK010931	BK010931 IL5-GN023
C 26	83.6	8.4	703	12 BM993909	BM993909 UI-H-DF1-
C 27	83.6	8.4	748	14 CA427898	CA427898 UI-H-DF0-
C 28	83.6	8.4	795	12 BG420907	BG420907 602452096
C 29	83.6	8.4	917	13 BX454417	BX454417 BX454417
C 30	83.6	8.4	1002	13 BX432291	BX432291 BX432291
C 31	83.4	8.4	367	10 BG014629	BG014629 IL5-GN023
C 32	83.2	8.4	222	9 AV738722	AV738722 AV738722
C 33	83.2	8.4	831	10 BF964558	BF964558 60267855
C 34	82.6	8.3	186	9 AI053535	AI053535 q172b05.x
C 35	82.6	8.3	190	9 AI053476	AI053476 q178h07.x
C 36	82.6	8.3	190	9 AI053526	AI053526 q172a07.x
C 37	82.6	8.3	190	9 AI053774	AI053774 q173h02.x
C 38	82.6	8.3	197	9 AI581244	AI581244 tb71b06.x
C 39	82.6	8.3	208	9 AI251122	AI251122 qv38h08.x
C 40	82.6	8.3	213	9 AI254486	AI254486 qv35b03.x
C 41	82.6	8.3	215	9 AI792510	AI792510 q172b05.y
C 42	82.6	8.3	221	9 AI345241	AI345241 tb81b11.x
C 43	82.6	8.3	225	9 AI792580	AI792580 q178c01.y
C 44	82.6	8.3	226	9 AI254649	AI254649 qv56h01.x
C 45	82.6	8.3	241	10 AW301538	AW301538 xs78d08.x

ALIGNMENTS

RESULT 1
BC039500
LOCUS BC039500
DEFINITION Homo sapiens, clone IMAGE:5555626, mRNA.
ACCESSION BC039500
VERSION BC039500.1 GI:25058499
KEYWORDS HTC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1457)
AUTHORS Strausberg,R.
TITLE Direct Submission
JOURNAL Submitted (01-NOV-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>
COMMENT Contact: MGC help desk
Email: cgabbs@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LINL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: <http://www.shgc.stanford.edu>

Contact: (Dickson, Mark) mcd@paxil.stanford.edu
 Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
 R. M.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LINL at: <http://image.llnl.gov>
 Series: IRAC Plate: 88 Row: g Column: 1
 This clone has the following problem: retained intron.

FEATURES

Location/Qualifiers
 1. 1467
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5555626"
 /tissue type="Uterus, leiomyosarcoma"
 /clone_lib="NIH MGC_71"
 /lab_host="DH10B"
 /note="Vector: PCMV-SPORT6"

ORIGIN

Query Match 91.4%; Score 904.6; DB 11; Length 1467;
 Best Local Similarity 98.8%; Pred. No. 3.1e-127;
 Matches 932; Conservative 0; Mismatches 9; Indels 2; Gaps 2;

QY 34 CGATCTCTGACCTGATCGCGCCGCTCAGCTCCCAAGTGTGGGATTACAGGCA 93
 DB 504 CGATCTCTGACCTGATCGCGCCGCTCAGCTCCCAAGTGTGGGATTACAGGCA 563
 QY 94 TGAGCACCGCGCGCGCCGCGACACTAGTTTAAAGGGCCCTCTGTGTGCGCGC 153
 DB 564 TGAGCACCGCGCGCGCCGCGACACTAGTTTAAAGGGCCCTCTGTGTGCGCGC 623
 QY 154 TGGCGCGCTCCAGCTGCCCCAGTCTGCGGGCTCAGTCCCGGTTCGCCATGTGTGGAG 213
 DB 624 TSCCGCGCTCCAGCTGCCCCAGTCTGCGGGCTCAGTCCCGGTTCGCCATGTGTGGAG 683
 QY 214 ACCGGTCCGTAAGCGCTGATGCTTCCGTGATGACATTTGGACCGGCTCTGGAC 273
 DB 684 ACCGGTCCGTAAGCGCTGATGCTTCCGTGATGACATTTGGACCGGCTCTGGAC 743
 QY 274 TGGGTAGGGAAGGAGGAGGCGGAATTTGGGCCCGAGGCGAGCTCTGCCACCCC 333
 DB 744 TGGGTAGGGAAGGAGGAGGCGGAATTTGGGCCCGAGGCGAGCTCTGCCACCCC 803
 QY 334 CGACTGGCGCTCCCGGTGGCCCCCGAGCGCTCCCGTGGCGCTTGGAGTCAAGTCTTAC 393
 DB 804 CGACTGGCGCTCCCGGTGGCCCCCGAGCGCTCCCGTGGCGCTTGGAGTCAAGTCTTAC 863
 QY 394 CGTCCGAGATCGTCCGCAACTGGCGAGCTGTGCATGGGCGGTGGCTAAGCGCGTGGTTT 453
 DB 864 CGTCCGAGATCGTCCGCAACTGGCGAGCTGTGCATGGGCGGTGGCTAAGCGCGTGGTTT 923
 QY 454 GGTACGATTGGCAGCGGAGCTTAAGTGTGTCTCTGAGAGCATGGACATTAGTCTGG 513
 DB 924 GGTACGATTGGCAGCGGAGCTTAAGTGTGTCTCTGAGAGCATGGACATTAGTCTGG 983
 QY 514 AGGGTCTCTGGAAGAGTGTATCCCGCCCCCACCATCAATGGCGTCTAGGTCTAGGAAGCGG 573
 DB 984 AGGGTCTCTGGAAGAGTGTATCCCGCCCCCACCATCAATGGCGTCTAGGTCTAGGAAGCGG 1043
 QY 574 GTGTGGGTGGGCGCTTAGGGCGAGGCGGAGACATACCCCGAAGTGGTGTGATTATACC 633
 DB 1044 GTGTGGGTGGGCGCTTAGGGCGAGGCGGAGACACACCCCGAAGTGGTGTGATTATACC 1103
 QY 634 GCAAGGGGCTGATTCGAACCCCCCAAGACACTGGAGGCTGTGTGCTGAGGAGGCGCC 693
 DB 1104 GCAAGGGGCTGATTCGAACCCCCCAAGACACTGGAGGCTGTGTGCTGAGGAGGCGCC 1163
 QY 694 GGCA-ATCCAGTGTGTGTGGGCTTTACAGGAAAGAGCTCCACCTTCT-TGGAGTGTGCA 751
 DB 1164 GGCAGATCCAGTGTGTGTGGGCTTTACAGGAAAGAGCTCCACCTTCTCTGGAGTGTGCA 1223
 QY 752 GATCGGATCTAGGTGTGTTCACCCGATGGGAGCTTGGGGCGGCGAGATGCTGCCCGCAGT 811

DB 1224 GATCGGATCTAGGTGTGTCCACCGATGGAGCTCGGGCGGCGAGATGCTGCCCGCAGT 1283
 QY 812 ACAAGAGCTGATTGGACCTGGGGCTCTGGATTCCTGATCTCTGCTGCTGATCTCCAG 871
 DB 1284 ACAAGAGCTGATTGGACCTGGGGCTCTGGATTCCTGATCTCTGCTGCTGATCTCCAG 1343
 QY 872 CAAAGTCTCTGCCCTGGCTTGCCTTCATCCACTCTCTCACTTCTCTGCCCTCAGAGTAA 931
 DB 1344 CAAAGTCTCTGCCCTGGCTTGCCTTCATCCACTCTCTCACTTCTCTGCCCTCAGAGTAA 1403
 QY 932 AATTGCAAGATCTGTGGTGCACAAAAAATAAAAAAATAAAAAA 974
 DB 1404 AATTGCAAGATCTGAAAAAATAAAAAAATAAAAAAATAAAAAA 1446

RESULT 2
 BX406138/c
 LOCUS
 DEFINITION
 BX406138 Homo sapiens FETAL LIVER Homo sapiens cDNA clone
 CS0DM010YD19 3-PRIME, mRNA sequence.
 ACCESSION
 VERSION
 BX406138
 BX406138.1 GI:30648317
 KEYWORDS
 EST.
 SOURCE
 Homo sapiens (human)
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
 Full-length cDNA libraries and normalization
 Unpublished (2001)
 CONTACT: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 EVRY cedex - France
 Email: secref@genoscope.cns.fr, Web: www.genoscope.cns.fr
 Library was constructed by Life Technologies, a division of
 Invitrogen. This sequence belongs to sequence cluster 8395.f For
 more information about this cluster, see
[http://www.genoscope.cns.fr/](http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CS0AM010CB10NP1&cluster=8395.f)
 cgi-bin/cluster.cgi?seq=CS0AM010CB10NP1&cluster=8395.f. Contact :
 Feng Liang Email: fliang@lifetech.com URL :
<http://fulllength.invitrogen.com/Invitrogen> Corporation 1600
 Faraday Avenue Genoscope sequence ID : CS0AM010CB10NP1.

FEATURES
 source
 1. 1201
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0DM010YD19"
 /tissue_type="FETAL LIVER"
 /dev_stage="fetal"
 /clone_lib="Homo sapiens FETAL LIVER"
 /note="Organ: liver; Vector: pCMVSPORT6; 1st strand cDNA
 was primed with a NotI-oligo(dt) primer. Five prime end
 enriched, double-strand cDNA was digested with Not I and
 cloned into the Not I and EcoRV sites of the pCMVSPORT 6
 vector. Library was not normalized."

ORIGIN

Query Match 80.2%; Score 793.8; DB 13; Length 1201;
 Best Local Similarity 94.9%; Pred. No. 1.7e-110;
 Matches 838; Conservative 29; Mismatches 11; Indels 5; Gaps 5;

QY 34 CGATCTCTGACCTGATCGCGCCGCTCAGCTCCCAAGTGTGGGATTACAGGCA 93
 DB 882 CGAKCKCKGACCTCTGATCCGCCCGCCCKAGCTCCCAAGGKGGGKACAGGCA 823
 QY 94 TGAGCACCGCGCGCGCCCGGACACTAGTATTTAAAGGCCCTCTGTGTGTCGCCG 153
 DB 822 TGAGCACCGCGCGCGCCCGGACACTAGTATTTAAAGGCCCTCTGTGTGTCGCCS 763
 QY 154 TGGCGCGCTCCAGCTGCCAGTCTGGGGGCTCAGTCCCGCGTTCCTATGTGTGGAG 213

Db	762	TGCGCGCGCTCCAGCTGCCAGCTTGGCGGGCTAGTCCCGGTTGCCATGTGGGAG	703
Qy	214	ACCGCGTCGCTAAGCGCTGGATGTGGCTTCGTGATGCACATTTGACCGGGCTCTGGAC	273
Db	702	ACCGCGTCGCGTAAAGCGCTGGATGTGGCTTCGTGATGCACATTTGACCGGGCTCTGGAC	643
Qy	274	TGGGCTAGGGGAAGGCGAGGCGGGAATTGGGCGGCGGAGGCGAGGCTCGCGACCCC	333
Db	642	TGGGCTAGGGGAAGGCGAGGCGGGAATT-GGCGGAGGCGGAGGCTCGCGACCCC	584
Qy	334	CGACTGCGCTCCCGGTGGCGCGCGCTCCCGGTGGCGCTCGAGTGCAGGCTTTAC	393
Db	583	CGACTGCGCTCCCGGTGGCGCGCGCTCCCGGTGGCGCTCGAGTGCAGGCTTTAC	524
Qy	394	CGTCCGAGATCGTCCGAACTGGCGGAGCTGTGCTGAGCGGCTTAAAGCGCTGGTT	453
Db	523	CGTCCGAGATCGTCCGAACTGGCGGAGCTGTGCTGAGCGGCTTAAAGCGCTGGTT	464
Qy	454	GGTTACGATTGGCGAGGCGGACTTAAGTGTGTCTCTGAAGACATGGACATTAGTCTGG	513
Db	463	GGTTACGATTGGCGAGGCGGACTTAAGTGTGTCTCTGAAGACATGGACATTAGTCTGG	404
Qy	514	AGGTCCTCGAAGATGATCCCGCCCACTCAATCAATAGCGCTTAGGCTTAGGAAGCGG	573
Db	403	AGGTCCTCGAAGATGATCCCGCCCACTCAATCAATAGCGCTTAGGCTTAGGAAGCGG	344
Qy	574	GTGTGGTGGGCGCTTAGGCGAGGCGGACACATACCCCAAGTGGTGGATTGTATACC	633
Db	343	GTGTGGTGGGCGCTTAGGCGAGGCGGACACATACCCCAAGTGGTGGATTGTATACC	284
Qy	634	GCAAGGGGCTGGATCGAAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	693
Db	283	GCAAGGGGCTGGATCGAAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	224
Qy	694	GGCA-ATCCAGTGTGCTGGGCTTTACAGGAAGAGCTCCACTTCT-TGGAGTGTGA	751
Db	223	GGCAGATCCAGTGTGCTGGGCTTTACAGGAAGAGCTCCACTTCT-TGGAGTGTGA	164
Qy	752	GATCGCATCTAGGTGTGCTCAACCGATGGGAGTG-CGGGCGGCGGAGATGCTGCCCGG	810
Db	163	GATCGCATCTAGGTGTGCTCAACCGATGGGAGTG-CGGGCGGCGGAGATGCTGCCCGG	104
Qy	811	TACAAAGCTGATTGGACCTGGGCGCTCTGGACCTTCCCTGATTCCTGCTTGCATCTCA	870
Db	103	TACAAAGCTGATTGGACCTGGGCGCTCTGGACCTTCCCTGATTCCTGCTTGCATCTCA	44
Qy	871	GCAAAGT-CCTGTCCTGGTGGCTGCTGCTTCACTTCACTCTCTCAC	912
Db	43	GMAAAGAACWGTGTCMGWTTGAAGCCATCATCCATCTCTCAC	1
RESULT 3			
BM980506/c			
LOCUS			
DEFINITION	UI-CF-EN1-add-a-08-0-UI.s1 UI-CF-EN1 Homo sapiens cDNA clone		
ACCESSION	BM980506		
VERSION	BM980506.1		
KEYWORDS	EST.		
SOURCE	Homo sapiens (human)		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
REFERENCE	1 (bases 1 to 760)		
AUTHORS	Bonald, M.F., Lennon, G. and Soares, M.B.		
TITLE	Normalization and subtraction: two approaches to facilitate gene discovery		
JOURNAL	Genome Res. 6 (9), 791-806 (1996)		
MEDLINE	97044477		
PUBMED	8889548		
COMMENT	Contact: McCray, PB		
	McCray Lab		

University of Iowa

2024 University of Iowa Med Labs, Iowa City, IA 52242, USA

Tel: 319 356 4866

Fax: 319 356 7171

Email: paul-mccray@uiowa.edu

Tissue Procurement: Dr. M. J. Welsh, University of Iowa

cDNA Library Arranged by: Dr. M. Bento Soares, University of Iowa

DNA Sequencing by: Dr. M. Bento Soares, University of Iowa

Clone Distribution: Researchers may obtain clones from Research Genetics (www.resgen.com) or from Open Biosystems (www.openbiosystems.com).

Seq primer: M13 FORWARD

POLYA=Yes

Location/Qualifiers

1..760

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="UI-CF-EN1-add-a-08-0-UI"

/tissue_type="Primary Lung Cystic Fibrosis Epithelial Cells"

/dev_stage="Adult"

/lab_host="DH10B (Life Technologies) (T1 phage resistant)"

/clone_lib="UI-CF-EN1"

/note="Organ: Lung; Vector: pT73-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I; UI-CF-EN1 is a normalized cDNA library containing the following tissue(s): Primary Lung Cystic Fibrosis Epithelial Cells. The library was constructed according to Bonald, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dr primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT73-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dr)18 tail. The sequence tag for this library is CTGCTCAGGT.

TAG TISSUE=Human Lung Epithelial Cell Lines untreated LPS 6hr to LPS 24h

TAG LIB=UI-CF-EN1

TAG_SEQ=CTGCTCAGGT"

ORIGIN

Query Match

Best Local Similarity

Matches

754;

Conservative

0;

Mismatches

4;

Indels

3;

Gaps

3;

72.6%;

99.1%;

754;

Score

719.2;

DB

12;

Length

760;

Qy

213

GACCGGTCGCTAAGCGCTGATGGCTGCTGATGCACATGGACCGGCTCTGGA

272

Db

760

GACCGGTCGCTAAGCGCTGATGGCTGCTGATGCACATGGACCGGCTCTGGA

702

Qy

273

CTGGGCTAGGGGAAGGCGAGGCGGGAATTTGGGCGGCGAGGCTCGCGACCC

332

Db

701

CTGGGCTAGGGGAAGGCGAGGCGGGAATTTGGGCGGCGAGGCTCGCGACCC

642

Qy

333

CCGACTGCGCTCCCGGTGGCGGCGGCGGCTCCCGGTGGCGGCTGAGTGCAGTCTTA

392

Db

641

CCGANTGCGCTCCCGGTGGCGGCGGCGGCTCCCGGTGGCGGCTGAGTGCAGTCTTA

582

Qy

393

CGGTCGAGATCGTCCGAACTGGCGGAGCTGTGCATGGGCGCTTAAAGCGCTGGTT

452

Db

581

CGGTCGAGATCGTCCGAACTGGCGGAGCTGTGCATGGGCGCTTAAAGCGCTGGTT

522

Qy

453

TGGTTACGATTGGCGGCGGCTTAAAGTGTGTCTCTGAAGAGCATGGACATTAGTCTG

512

Db

521

TGGTTACGATTGGCGGCGGCTTAAAGTGTGTCTCTGAAGAGCATGGACATTAGTCTG

462

Qy

513

GAGGGTCTTGGAGAGTGTATCCCGGCGGCGGCGGCTTAGTCTAGGAAGCG

572

Db

461

GAGGGTCTTGGAGAGTGTATCCCGGCGGCGGCTTAGTCTAGGAAGCG

402

573 GGTGGTGGGGCTTAGGGCGAGCGCAGACATACCCCGAAGTGGTGGATTGTATAC 632
 401 GGTGGTGGGGCTTAGGGCGAGCGCAGACACACCCCGAAGTGGTGGATTGTATAC 342
 633 CGCAAGGGCTGGATCGAACCCCAAGACACTGGAAGGCTGTGTGGCTGAGAGGGCC 692
 341 CGCAAGGGCTGGATCGAACCCCAAGACACTGGAAGGCTGTGTGGCTGAGAGGGCC 282
 693 CGGCA-ATCCAGTGTGTGGCTTTACAGGAAGAGCTCCACCTTCT-TGGAGTGTGC 750
 281 CGGAGATCCAGTGTGTGGCTTTACAGGAAGAGCTCCACCTTCT-TGGAGTGTGC 222
 751 AGATGCGATCTAGTGTGTGTCCACCGATGGAGCTGGGGCGGCGAGATGCTGCCCCAG 810
 221 AGATGCGATCTAGTGTGTGTCCACCGATGGAGCTGGGGCGGCGAGATGCTGCCCCAG 162
 811 TACAAAGCTGATTGGACCTGGGGCTCTGACCTTCCCTGATCTCTGTGATCTTCCA 870
 161 TACAAAGCTGATTGGACCTGGGGCTCTGACCTTCCCTGATCTCTGTGATCTTCCA 102
 871 GCAAAGTCTCTCCGCTGGCTGCTTCATCCACTCTCTCACTTCTCTGCTTTCAGAGTA 930
 101 GCAAAGTCTCTCCGCTGGCTGCTTCATCCACTCTCTCACTTCTCTGCTTTCAGAGTA 42
 931 AAATGCAAGATCTGTGGTGCACCAAAAAAAAAAAAAAAAAAAAA 971
 41 AAATGCAAGATCTGTGGTGCCTTAAAAAAAAAAAAAAAAAAAA 1

RESULT 4
 BU620427/c
 LOCUS
 DEFINITION UI-H-FL1-bfw-n-19-0-UI.s1 NCI CGAP FL1 Homo sapiens cDNA clone
 UI-H-FL1-bfw-n-19-0-UI 3', mRNA sequence.
 ACCESSION BU620427
 VERSION BU620427.1 GI:23286642
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 691)
 NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 Unpublished (1997)
 Contact: Robert Strausberg, Ph.D.
 Email: cgaps-remail.nih.gov
 Tissue Procurement: James Martin
 cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Clone distribution information can be obtained
 from Dr. M. Bento Soares, bento-soares@uiowa.edu
 Seq primer: M13 FORWARD
 POLYA=Yes

Location/Qualifiers
 1..691
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="UI-H-FL1-bfw-n-19-0-UI"
 /tissue_type="Cell lines"
 /dev_stage="Adult"
 /lab_host="DR10B (Life Technologies)"
 /clone_lib="NCI CGAP FL1"
 /note="Organ: Chondrosarcoma; Vector: pT73-Pac
 (Pharmacia) with a modified polylinker; Site 1: EcoR I;
 Site 2: Not I; NCI CGAP FL1 is a normalized cDNA library
 derived from a pool of mRNA obtained from 4 cell lines
 from grade III chondrosarcoma tissues. The library was
 constructed according to Bonaldo, Lennon and Soares,
 Genome Research, 6:791-806, 1996. First strand cDNA

FEATURES
 source

synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT73-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is GAGTCCGGTG. The cell lines were provided by Dr. James Martin from the University of Iowa.
 TAG_RISSE=Human Chondrosarcoma Grade 3 cell line mix
 TAG_LI=UI-H-FL1
 TAG_SEQ=GAGTCCGGTG"

ORIGIN

Query Match 65.8%; Score 651.2; DB 13; Length 691;
 Best Local Similarity 99.1%; Pred. No. 6.9e-89;
 Matches 666; Conservative 0; Mismatches 3; Indels 3; Gaps 3;

QY 279 TAGGGGAAGGCGAGGAGGCGGAATTGGGCCCGAGGCGCAGGCTCGCGACCCCGGACT 338
 DB 691 TAGGGGAAGGCGAGGAGGCGGAA-TGGGCCCGAGGCGCAGGCTCGCGACCCCGGACT 633
 QY 339 GGGCTCCGGTGGCCCGGAGCGCTCCCGGTGGCCCTGGAGTGCAGGTCTTACCGTCC 398
 DB 632 GCGCTCCGGTGGCCCGGAGCGCTCCCGGTGGCCCTGGAGTGCAGGTCTTACCGTCC 573
 QY 399 GAGATCGTCCGCACTGGGCGGAGCTGTGCATGGGCGGTGGCTAAGGCGCGTGGTTGGTTA 458
 DB 572 GAGATCGTCCGCACTGGGCGGAGCTGTGCATGGGCGGTGGCTAAGGCGCGTGGTTGGTTA 513
 QY 459 CGATTGGCGAGCGGACTTAAGTGTGTCTCTGAAGAGCATGGAACATTAGTCTCGAGGGT 518
 DB 512 CGATTGGCGAGCGGACTTAAGTGTGTCTCTGAAGAGCATGGAACATTAGTCTCGAGGGT 453
 QY 519 CTTGAAGAGTGTATCCCGCCCGCCACCATCAATCGCGCTTAGGTCTAGGAAGCGGGTGTG 578
 DB 452 CTTGAAGAGTGTATCCCGCCCGCCACCATCAATCGCGCTTAGGTCTAGGAAGCGGGTGTG 393
 QY 579 GGTGGGGCTTTAGGCGGAGGCGCAGACATACCCGAAAGTGGTGGATTGTATACCGCAAG 638
 DB 392 GGTGGGGCTTTAGGCGGAGGCGCAGACACACCCGAAAGTGGTGGATTGTATACCGCAAG 333
 QY 639 GGGCTGGATCGAACCCCGCCCGGAGCAGTGGAGCTGTGGCTGAGGAGGCGCGGCA- 697
 DB 332 GGGCTGGATCGAACCCCGCCCGGAGCAGTGGAGCTGTGGCTGAGGAGGCGCGGCGAG 273
 QY 698 ATCCAGTGTGTGGGGCTTTACAGGAAGAGCTCCACCTTCT-TGGAGTGTGCAGATGC 756
 DB 272 ATCCAGTGTGTGGGGCTTTACAGGAAGAGCTCCACCTTCTCTGGAGTGTGCAGATGC 213
 QY 757 GATCTAGTGTGTCCACCCGATGGAGCTGGGGCGGCGGAGATGCTGCCCGCAGTACAAA 816
 DB 212 GATCTAGTGTGTCCACCCGATGGAGCTGGGGCGGCGGAGATGCTGCCCGCAGTACAAA 153
 QY 817 GCTGATTTGGACCTGGGGCTCTGGACTTCCCTGATCTCTGCTTGCATCTCCAGCAAG 876
 DB 152 GCTGATTTGGACCTGGGGCTCTGGACTTCCCTGATCTCTGCTTGCATCTCCAGCAAG 93
 QY 877 TCCGTGCCGTGGCTGGCTTCATCCACTCTCTCACTTCTGCTTTCAGAGTAAATTTG 936
 DB 92 TCCGTGCCGTGGCTGGCTTCATCCACTCTCTCACTTCTGCTTTCAGAGTAAATTTG 33
 QY 937 CAAGATCTGTGTGCACCAAAAAAAAAAAAAAAAAAAAA 968
 DB 32 CAAGATCTGTGTGCACCAAAAAAAAAAAAAAAAAAAAA 1

RESULT 5
 BU684594/c
 LOCUS
 DEFINITION UI-CF-EN1-act-a-22-0-UI.s1 UI-CF-EN1 Homo sapiens cDNA clone
 UI-CF-EN1-act-a-22-0-UI 3', mRNA sequence.
 ACCESSION BU684594

VERSION	BU684594.1	GI:23537704	
KEYWORDS	EST.		
SOURCE	Homo sapiens	(human)	
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	1 (bases 1 to 694)		
TITLE	Bonaldo,M.F., Lennon,G. and Soares,M.B.		
JOURNAL	Normalization and subtraction: two approaches to facilitate gene		
MEDLINE	discovery		
PUBMED	Genome Res. 6 (9), 791-806 (1996)		
COMMENT	97044477		
	8889548		
	Contact: McCray, PB		
	McCray Lab		
	University of Iowa		
	2024 University of Iowa Med Labs, Iowa City, IA 52242, USA		
	Tel: 319 356 4866		
	Fax: 319 356 7171		
	Email: paul-mccray@uiowa.edu		
	Tissue Procurement: Dr. M. J. Welsh, University of Iowa		
	cDNA Library preparation: Dr. M. Bento Soares, University of Iowa		
	cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa		
	DNA Sequencing by: Dr. M. Bento Soares, University of Iowa		
	Clone Distribution: Researchers may obtain clones from Research		
	Genetics (www.resgen.com) or from Open Biosystems		
	(www.openbiosystems.com).		
	Seq primer: M13 FORWARD		
	POLYA=Yes		
FEATURES	Location/Qualifiers		
source	1..694		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	/clone="UI-CF-EN1-act-a-22-0-UI"		
	/tissue_type="Primary Lung Cystic Fibrosis Epithelial		
	Cells"		
	/dev_stage="Adult"		
	/lab_host="D310B (Life Technologies) (T1 phage resistant)"		
	/clone_lib="UI-CF-EN1"		
	/note="Organ: Lung; Vector: pT73-Pac (Pharmacia) with a		
	modified polylinker; Site 1: EcoR I; Site 2: Not I;		
	UI-CF-EN1 is a normalized cDNA library containing the		
	following tissue(s): Primary Lung Cystic Fibrosis		
	Epithelial Cells. The library was constructed according to		
	Bonaldo, Lennon and Soares, Genome Research, 6:791-806,		
	1996. First strand cDNA synthesis was primed with an		
	oligo-dT primer containing a Not I site. Double stranded		
	cDNA was ligated to an EcoR I adaptor, digested with Not		
	I, and cloned directionally into pT73-Pac vector. The		
	oligonucleotide used to prime the synthesis of		
	first-strand cDNA contains a library tag sequence that is		
	located between the Not I site and the (dT)18 tail. The		
	sequence tag for this library is CTGCTCAGGT.		
	TAG TISSUE=Human Lung Epithelial Cell Lines untreated LPS		
	6hr to LPS 24h		
	TAG_LIB=UI-CF-EN1		
	TAG_SEQ=CTGCTCAGGT"		
ORIGIN			
Query Match	63.7%; Score 630.8; DB 13; Length 694;		
Best Local Similarity	97.7%; Pred. No. 8.3e-86;		
Matches	671; Conservative 0; Mismatches 13; Indels 3; Gaps 3;		
QY	290 CAGAGGGCGGAATGGCGCGAGGGCCAGGCTTGGCCGACCCCGAGTCTGCGCTCCCGG 349		
Db	694 CAGGAGGGCGGAA-TGGGCGCGAGGGCCAGGCTTGGCCGACCCCGAGTCTGCGCTCCCGG 636		
QY	350 TGGCCCCGAGCGCTCCCGGTGGCTGAGTGCGAGTCTTACCGTCCGAGATCTCCG 409		
Db	635 TGGCCCCGAGCGCTCCCGGTGGCTGAGTGCGAGTCTTACCGTCCGAGATCTCCG 576		
QY	410 CAACCTGGCGAGCTGTGCATGGGCGGTGGCTTAAGGCGGTGGTTGGTTACGATTGGCCAG 469		

Db	575	CAACTGGGGGAGCTGTGCATGGGGCGTGGCTAAGCCCGTGGTTGGTTACGATTGGCCAG	516
QY	470	CGGGACTTAAGTGTGTCTCTGAAGAGCATGAGACATTAGTCTGGAGGTCCTGGAAGAGT	529
Db	515	CGGGACTTAAGTGTGTCTCTGAAGAGCATGAGACATTAGTCTGGAGGTCCTGGAAGAGT	456
QY	530	GATCCCCCGCCACCATCAATGGCGCTTAGTCTAGGAAGCGGTGGGTGGGGCCCTT	589
Db	455	GATCCCCCGCCACCATCAATGGCGCTTAGTCTAGGAAGCGGTGGGTGGGGCCCTT	396
QY	590	AGGGCGAGGCGCAGACATACCCCGAAGTGGTTGGATTATACCGCAAGGGGCTGGATCG	649
Db	395	AGGGCGAGGCGCAGACATACCCCGAAGTGGTTGGATTATACCGCAAGGGGCTGGATCG	336
QY	650	AACCCCCCAAGACACTGGAGGCTGTGTGGCTGAGGAGGCGCCCGCA-ATCCAGTGTGT	708
Db	335	AACCCCCCAAGACACTGGAGGCTGTGTGGCTGAGGAGGCGCCCGCAATCCAGTGTGT	276
QY	709	CGTGGGCTTTACAGAAAGAGCTCCACCTTCT-TGGAGTGTGCAGATCGCATCTAGGTGT	767
Db	275	CGTGGGCTTTACAGAAAGAGCTCCACCTTCTTCTGGAGTGTGCAGATCGCATCTAGGTGT	216
QY	768	GTCCACCCGATGGGAGCTGGGGCGGGCAGATGCTGCCCGCAGTACAAAGCTGATTGGA	827
Db	215	GTCCACCCGATGGGAGCTGGGGCGGGCAGATGCTGCCCGCAGTACAAAGCTGATTGGA	156
QY	828	CCTGGGCGCTCTGGACTTCCCTGATTCCTGCTTCATCTCCAGCAAGTCTCTGCCGT	887
Db	155	CCTGGGCGCTCTGGACTTCCCTGATTCCTGCTTCATCTCCAGCAAGTCTCTGCCGT	96
QY	888	TGGCTGCCCTTCATCCACTCTCTCACTTCTGCTTCCGCTTCAGAGTAAATTCGAGATCTGTG	947
Db	95	TGGCTGCCCTTCATCCACTCTCTCACTTCTGCTTCCGCTTCAGAGTAAATTCGAGATCTGTG	36
QY	948	GTGCAAAAAAAAAAAAAAAAAAAAAAAAAA 974	
Db	35	GTGCTTACTGGGATCTGAATAAAAAAAAA 9	
RESULT	6		
LOCUS	BM679577/1		
DEFINITION	UI-E-EOO-aia-1-05-0-UI.s1 UI-E-EOO Homo sapiens cDNA clone		
ACCESSION	BM679577		
VERSION	BM679577.1		
KEYWORDS	EST.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	1 (bases 1 to 669)		
AUTHORS	Bonaldo,M.F., Lennon,G. and Soares,M.B.		
TITLE	Normalization and subtraction: two approaches to facilitate gene		
JOURNAL	discovery		
MEDLINE	Genome Res. 6 (9), 791-806 (1996)		
PUBMED	97044477		
COMMENT	8889548		
	Contact: Soares, MB		
	Coordinated Laboratory for Computational Genomics		
	University of Iowa		
	375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA		
	Tel: 319 335 8250		
	Fax: 319 335 9565		
	Email: bento-soares@uiowa.edu		
	Tissue Procurement: Dr. Gregg Hageman		
	cDNA Library preparation: Dr. M. Bento Soares, University of Iowa		
	cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa		
	DNA Sequencing by: Dr. M. Bento Soares, University of Iowa		
	Clone Distribution: Researchers may obtain clones from Research		
	Genetics (www.resgen.com).		
	Seq primer: M13 Forward		

```
FEATURES          Location/Qualifiers
  source
1..669
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-E-E00-aia-1-05-0-UI"
/tissue_type="fetal eye"
/dev_stage="fetal"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/clone_lib="UI-E-E00"
/note="Organ: eye; Vector: pT7T3-Pac (Pharmacia) with a
modified polylinker; Site 1: EcoR I; Site 2: Not I;
UI-E-E00 is a cDNA library containing the following
tissue(s): fetal eye. The library was constructed
according to Bonaldo, Lennon and Soares, Genome Research,
6:791-806, 1996. First strand cDNA synthesis was primed
with an oligo-dT primer containing a Not I site. Double
stranded cDNA was ligated to an EcoR I adaptor, digested
with Not I, and cloned directionally into pT7T3-Pac
vector. The oligonucleotide used to prime the synthesis of
first-strand cDNA contains a library tag sequence that is
located between the Not I site and the (dT)18 tail. The
sequence tag for this library is CGGTATACC. This library
System, supported by National Eye Institute (NEI).
TAG_TISSUE=human fetal eye
TAG_LIB=UI-E-E00
TAG_SEQ=CGGTATACC"

ORIGIN
Query Match      63.5%; Score 628.2; DB 12; Length 669;
Best Local Similarity 98.3%; Pred. No. 2.1e-85;
Matches 655; Conservative 0; Mismatches 9; Indels 2; Gaps 2;

QY 311 GAGGCCAGGCTCGCCGACCCCGACTGCGCTCCCGTGGGCGCCCGCAGCGCTCCCGG 370
DB 669 GAGGCCAGGCTCGCCGACCCCGANTGCGCTCCCGTGGGCGCCCGCAGCGCTCCCGG 610
QY 371 TGGCCCTGGAGTCAGGTCTTACCGTCCGAGATCGTCCGCAACTGGCGGAGCTGTGCATG 430
DB 609 TGGCCCTGGAGTCAGGTCTTACCGTCCGAGATCGTCCGCAACTGGCGGAGCTGTGCATG 550
QY 431 GGGCGTGGCTAAGCGCGTGGTTGGTTACGATTGGCCAGCGGACCTTAAGTGTGTCTCT 490
DB 549 GGGCGTGGCTAAGCGCGTGGTTGGTTACGATTGGCCAGCGGACCTTAAGTGTGTCTCT 490
QY 491 GAAGAGATGACATTAGTCTGGAGGGTCTTGGGAAGAGTATCCCGCCCGCCACCATCAA 550
DB 489 GAAGAGATGACATTAGTCTGGAGGGTCTTGGGAAGAGTATCCCGCCCGCCACCATCAA 430
QY 551 TGGCGCTTAGGTCTAGGAAGCGGTGTGGTGGGGCTTAGGGCGAGGCGCAGCATACC 610
DB 429 TGGCGCTTAGGTCTAGGAAGCGGTGTGGTGGGGCTTAGGGCGAGGCGCAGCATACC 370
QY 611 CCGAAGTGGTGGATTGTATACCGCAAGGGGCTGGATCGAACCCCGCCAAAGACACTGGA 670
DB 369 CCGAAGTGGTGGATTGTATACCGCAAGGGGCTGGATCGAACCCCGCCAAAGACACTGGA 310
QY 671 GGCTGTGTGGCTAGGAGGGGCCCGGCA-ATCCAGTGTGTCTGGGCTTTACAGAAAGAG 729
DB 309 GGCTGTGTGGCTAGGAGGGGCCCGGCAAGTGTGTCTGGGCTTTACAGAAAGAG 250
QY 730 CTCACCTTCT-TGGAGTGCAGATCGATCTAGTGTGTCTCCACCGCATGGAGCTGG 788
DB 249 CTCACCTTCT-TGGAGTGCAGATCGATCTAGTGTGTCTCCACCGCATGGAGCTGG 190
QY 789 GCGCGGGCAGATGCTGCCCGCAGTACAAAGCTGATTTCGACCTGGGGCTCTGGACCTCCC 848
DB 189 GCGCGGGCAGATGCTGCCCGCAGTACAAAGCTGATTTCGACCTGGGGCTCTGGACCTCCC 130
QY 849 TGATTCTCTGTTGCATCTCCAGCAAGTCTGTCCCGTTGGCTGCCTTCATCCACTCTC 908
```

```
DB 129 TGATTCTCTGCTTGCATCTCCAGCAAGTCTGCTCCGTTGGCTGCCTTCATCCACTCTC 70
QY 909 TCATTCTCTGCTTGCATCTCCAGTAAATTCAGATCTGTGGTGCAGAAAAA 968
DB 69 TCATTCTCTGCTTGCATCTCCAGTAAATTCAGATCTGTGGTGCAGAAAAA 10
QY 969 AAAAAA 974
DB 9 AAAAAA 4

RESULT 7
BU617331/c
LOCUS
DEFINITION
632 bp mRNA linear EST 23-SEP-2002
UI-H-DF0-bep-n-09-0-UI.s1 NCI CGAP DF0 Homo sapiens cDNA clone
UI-H-DF0-bep-n-09-0-UI 3', mRNA sequence.
ACCESSION
BU617331
VERSION
BU617331.1 GI:23283539
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1. (bases 1 to 632)
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
AUTHORS
Tumor Gene Index
TITLE
Unpublished (1997)
JOURNAL
Contact: Robert Strausberg, Ph.D.
COMMENT
Email: cgapbs-remail.nih.gov
Tissue Procurement: Dr. Jose Mercuende
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone distribution information can be obtained
from Dr. M. Bento Soares, bento-soares@uiowa.edu
Seq primer: M13 FORWARD
POLYA=Yes.

FEATURES
Location/Qualifiers
1..632
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-H-DF0-bep-n-09-0-UI"
/tissue_type="Subchondral Bone"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies)"
/clone_lib="NCI CGAP DF0"
/note="Organ: Bone; Vector: pT7T3-Pac (Pharmacia) with a
modified polylinker; Site 1: EcoR I; Site 2: Not I;
NCI CGAP DF0 is a cDNA library containing the following
tissue(s): Subchondral Bone. The library was constructed
according to Bonaldo, Lennon and Soares, Genome Research,
6:791-806, 1996. First strand cDNA synthesis was primed
with an oligo-dT primer containing a Not I site. Double
stranded cDNA was ligated to an EcoR I adaptor, digested
with Not I, and cloned directionally into pT7T3-Pac
vector. The oligonucleotide used to prime the synthesis of
first-strand cDNA contains a library tag sequence that is
located between the Not I site and the (dT)18 tail. The
sequence tag for this library is GTTAAGGCTC.
TAG_TISSUE=Subchondral bone
TAG_LIB=UI-H-DF0
TAG_SEQ=GTTAAGGCTC"

ORIGIN
Query Match      60.9%; Score 603.2; DB 13; Length 632;
Best Local Similarity 99.2%; Pred. No. 1.2e-81;
Matches 627; Conservative 0; Mismatches 3; Indels 2; Gaps 2;

QY 339 GGGCTCTCCGGTGGCCCGCAGCGCTCCCGTGGCGTGGAGTGTACCGTCC 398
DB 632 GGGCTCTCCGGTGGCCCGCAGCGCTCCCGTGGCGTGGAGTGTACCGTCC 573
```


QY 399 GAGATCGTCGCGCACTGGCGAGCTGTGCATGGGCGTGCTAAGGCCGTGGTTGGTTA 458
 Db |||||
 QY 572 GAGATCGTCGCGCACTGGCGAGCTGTGCATGGGCGTGCTAAGGCCGTGGTTGGTTA 513
 Db |||||
 QY 459 CGATTGCCAGCGGAGCTTAAGTGTCTCTCTGAAGAGCATGGACATTAGTCTGGAGGGT 518
 Db |||||
 QY 512 CGATTGCCAGCGGAGCTTAAGTGTCTCTCTGAAGAGCATGGACATTAGTCTGGAGGGT 453
 Db |||||
 QY 519 CTTGGAGAGTGATCCCGGCCACCATCAATCAATGGCGCTTAGTCTAGGAAGCGGGTGTG 578
 Db |||||
 QY 452 CTTGGAGAGTGATCCCGGCCACCATCAATCAATGGCGCTTAGTCTAGGAAGCGGGTGTG 393
 Db |||||
 QY 579 GGTGGGCGCTTAGGGCGAGCGGAGACATACCCCGAGTGGTGGATTGTATACCGCAAG 638
 Db |||||
 QY 392 GGTGGGCGCTTAGGGCGAGCGGAGACACACCCCGAGTGGTGGATTGTATACCGCAAG 333
 Db |||||
 QY 639 GGGCTGATCGAACCCTCCCAAGACATGGAAGGCTGTGTGGCTGAGGAGGGCCCGGCA- 697
 Db |||||
 QY 332 GGGCTGATCGAACCCTCCCAAGACATGGAAGGCTGTGTGGCTGAGGAGGGCCCGGCA 273
 Db |||||
 QY 698 ATCCAGTGTGTCTGGGCTTTACAGGAAGAGCTCCACCTTC- TGGAGTGTGCAGATGC 756
 Db |||||
 QY 272 ATCCAGTGTGTCTGGGCTTTACAGGAAGAGCTCCACCTTC- TGGAGTGTGCAGATGC 213
 Db |||||
 QY 757 GATCTAGTGTGTCTACCCGATGGAGCTGCGGCGGGGAGATGCTGCCCGCAGTACAAA 816
 Db |||||
 QY 212 GATCTAGTGTGTCTACCCGATGGAGCTGCGGCGGGGAGATGCTGCCCGCAGTACAAA 153
 Db |||||
 QY 817 GCTGATTTGACCTGGGCGCTTGGACTTCCCTGATCTCTGCTGCACTCCAGCAAG 876
 Db |||||
 QY 152 GCTGATTTGACCTGGGCGCTTGGACTTCCCTGATCTCTGCTGCACTCCAGCAAG 93
 Db |||||
 QY 877 TCCTGTCCCGTGGCTGCCCTTCATCCACTCTCTCACTTCTCTGCTTCAGAGTAAATTTG 936
 Db |||||
 QY 92 TCCTGTCCCGTGGCTGCCCTTCATCCACTCTCTCACTTCTCTGCTTCAGAGTAAATTTG 33
 Db |||||
 QY 937 CAGATCTGTGGTGCAGAAAAA 968
 Db |||||
 QY 32 CAAGATCTGTGGTGCAGAAAAA 1

RESULT 8
 CA313142/c
 LOCUS
 DEFINITION UI-CF-FNO-aex-n-23-0-UI.s1 UI-CF-FNO Homo sapiens cDNA clone
 ACCESSION CA313142
 VERSION CA313142.1 GI:24531240
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 633)
 AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.
 TITLE Normalization and subtraction: two approaches to facilitate gene
 discovery
 JOURNAL Genome Res. 6 (9), 791-806 (1996)
 MEDLINE 97044477
 PUBMED 8889548
 COMMENT Contact: McCray, PB
 University of Iowa
 2024 University of Iowa Med Labs, Iowa City, IA 52242, USA
 Tel: 319 356 4866
 Fax: 319 356 7171
 Email: paul-mccray@uiowa.edu
 Tissue Procurement: Dr. M. J. Welsh, University of Iowa
 CDNA Library preparation: Dr. M. Bento Soares, University of Iowa
 CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Researchers may obtain clones from Research

Genetics (www.reagen.com) or from Open Biosystems
 (www.openbiosystems.com).
 Seq primer: M13 FORWARD
 POLYA=Yes.

FEATURES

source
 Location/Qualifiers
 1..633
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="UI-CF-FNO-aex-n-23-0-UI"
 /tissue_type="Human Lung Epithelial cells"
 /lab_host="DH10B (Life Technologies) (T1 phage resistant)"
 /clone_lib="UI-CF-FNO"
 /note="Organ: Lung; Vector: pT73-Pac (Pharmacia) with a
 modified polylinker; Site 1: EcoR I; Site 2: Not I; two
 UI-CF-FNO is a subtracted cDNA library derived from
 normalized Human lung epithelial cell libraries (ENI and
 DU1) The library was subtracted according to according to
 Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
 1996. For additional information, contact:
 bento-soares@uiowa.edu
 TAG_TISSUE=Human Lung Epithelial Cell Lines untreated LPS
 shr to LPS 24h
 TAG_L1B=UI-CF-FNO
 TAG_SEQ=CTGCTCAGGT"

ORIGIN

Query Match 60.9%; Score 602.6; DB 14; Length 633;
 Best Local Similarity 99.1%; Pred. No. 1.5e-81;
 Matches 627; Conservative 0; Mismatches 4; Indels 2; Gaps 2;
 QY 341 GCCTCCCGGTGGCCCGCAGCGCCTCCCGGTGGCCCTCGAGTGCAGGCTTCTTACCGTCCGA 400
 Db 633 GCCTCCCGGTGGCCCGCAGCGCCTCCCGGTGGCCCTCGAGTGCAGGCTTCTTACCGTCCGA 574
 QY 401 GATCTCGCAACTGGCGAGCTGTGCATGGGCGTGTGCTAAGCCGTGGTTGGTTAGC 460
 Db 573 GATCTCGCAACTGGCGAGCTGTGCATGGGCGTGTGCTAAGCCGTGGTTGGTTAGC 514
 QY 461 ATTGGCCAGCGGACTTAAGTGTGTCTCTCAAGAGCATGGACATTAGTCTGGAGGGTCC 520
 Db 513 ATTGGCCAGCGGACTTAAGTGTGTCTCTCAAGAGCATGGACATTAGTCTGGAGGGTCC 454
 QY 521 TGAAGAGTGATCCCGCCCAACCATTAATGGCGCTTAGGTCTAGGAAGCGGGTGTGGG 580
 Db 453 TGAAGAGTGATCCCGCCCAACCATTAATGGCGCTTAGGTCTAGGAAGCGGGTGTGGG 394
 QY 581 TGGGCGCTTAGGGCGAGCGGAGCATACCCGAGTGGTTGGATTGTATACCGCAAGGG 640
 Db 393 TGGGCGCTTAGGGCGAGCGGAGCATACCCGAGTGGTTGGATTGTATACCGCAAGGG 334
 QY 641 GGTGATCGAAACCCCGCAAGACATGGAAGGCTGTGTGGCTGAGGAGGGCCCGGCA-AT 699
 Db 333 GGTGATCGAAACCCCGCAAGACATGGAAGGCTGTGTGGCTGAGGAGGGCCCGGCAAT 274
 QY 700 CCAGTGTGTGGCTTTACAGGAAGAGCTCCACCTTC- TTGAGTGTGCAGATGCGA 758
 Db 273 CCAGTGTGTGGCTTTACAGGAAGAGCTCCACCTTC- TTGAGTGTGCAGATGCGA 214
 QY 759 TCTAGGTGTCTCCACCCGATGGAGCTGCGGCGGGGAGATGTGCTGCCCGAGTACAAAGC 818
 Db 213 TCTAGGTGTCTCCACCCGATGGAGCTGCGGCGGGGAGATGTGCTGCCCGAGTACAAAGC 154
 QY 819 TGATTGACCTGGGGCCTCTGGACTTCCCTGATTCTCTGTTGATCTCCAGCAAGTGC 878
 Db 153 TGATTGACCTGGGGCCTCTGGACTTCCCTGATTCTCTGTTGATCTCCAGCAAGTGC 94
 QY 879 CTGTCCCGTGTGGCTTCATCCACTCTCCACTTCTCGCTTCAGAGTAAATTTGCA 938
 Db 93 CTGTCCCGTGTGGCTTCATCCACTCTCCACTTCTCGCTTCAGAGTAAATTTGCA 34
 QY 939 AGATCTGTGGTGCAGAAAAA 971
 Db |||||

```

Db      33 AGATCTGGTGTCTTAAAAAATAAAAAAAAAAAAAA 1
        CD742581      623 bp      mRNA      linear      EST 26-JUN-2003
        UI-H-FT2-bj1-i-10-0-UI.s1 NCI CGAP_FT2 Homo sapiens cDNA clone
        UI-H-FT2-bj1-i-10-0-UI 3', mRNA sequence.

RESULT 9
CD742581/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 623)
NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Dr. Gary W. Hunninghake, U of I
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at
http://genome.uiowa.edu/distribution/cgap.html
Seq primer: M13 FORWARD
POLYA=yes.

FEATURES
source
1..623
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-H-FT2-bj1-i-10-0-UI"
/tissue_type="Aveolar Macrophage"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies)"
/clone_lib="NCI CGAP FT2"
/note="Organ: Lung; Vector: p7T3-Pac (Pharmacia) with a
modified polylinker; Site 1: EcoR I; Site 2: Not I;
NCI-CCAP FT2 is a subtracted cDNA library constructed from
a pool of 81 RNA samples from Alveolar Macrophages
challenged with different treatments. The library was
subtracted according to Bonaldo, Lennon and Soares, Genome
Research, 6:791-806, 1996. The tissue was provided by Dr.
Gary W. Hunninghake of the University of Iowa."
TAG_TISSUE=Human Lung Aveolar Macrophage
TAG_LIB=UI-H-FT2
TAG_SEQ=GGCCATGCGC

ORIGIN
Query Match 59.4%; Score 587.8; DB 14; Length 623;
Best Local Similarity 98.6%; Pred. No. 2.6e-79;
Matches 614; Conservative 0; Mismatches 7; Indels 2; Gaps 2;

Qy      343 CTCGGGTGGCCCGCAGCGCCTCCCGTGGCCCTGAGTGCAGGTCTTACCGTCCGAGA 402
Db      623 CTCGGGTGGCCCGCAGCGCCTCCCGTGGCCCTGAGTGCAGGTCTTACCGTCCGAGA 564
Qy      403 TCGTCCCACTGGCGAGCTGTCATGGGCGTGGCTAAGGCGGTGGTTGGTTAGCAT 462
Db      563 TCGTCCCACTGGCGAGCTGTCATGGGCGTGGCTAAGGCGGTGGTTGGTTAGCAT 504
Qy      463 TGGCCAGCGGACTTAAGTGTGTCTCTGAAGAGCATGACATTAGTCTGGAGGCTCTG 522
Db      503 TGGCCAGCGGACTTAAGTGTGTCTCTGAAGAGCATGACATTAGTCTGGAGGCTCTG 444
Qy      523 GAAGAGTATCCCGCCCGCCACCAATGCGCTTAGGTCTAGGAGCGGGTCTGGTG 582
Db      443 GAAGAGTATCCCGCCCGCCCGCCACCAATGCGCTTAGGTCTAGGAGCGGGTCTGGTG 384

```

```

Qy      583 GGGCTTAGGGCGAGCGCAGACATACCCGAGTGGTTGGATTGTATACCGAAGGGC 642
Db      383 GGGCTTAGGGCGAGCGCAGACATACCCGAGTGGTTGGATTGTATACCGAAGGGC 324
Qy      643 TGGATCGAACCCCAAGACACTGGAAGGCTGTGTGCTTGGAGGGCCCGGCA-ATCC 701
Db      323 TGGATCGAACCCCAAGACACTGGAAGGCTGTGTGCTTGGAGGGCCCGGCAAGATCC 264
Qy      702 AGTGTCTGTGGGCTTTACAGGAAGAGCTCCACCTTCT-TGGAGTGTGCAGATCGCATC 760
Db      263 AGTGTCTGTGGGCTTTACAGGAAGAGCTCCACCTTCT-TGGAGTGTGCAGATCGCATC 204
Qy      761 TAGGTGTCTCCACCCGATGGGAGCTGGGGCGGGCAGATGCTGCCCCAGTACAAAGCTG 820
Db      203 TAGGTGTCTCCACCCGATGGGAGCTGGGGCGGGCAGATGCTGCCCCAGTACAAAGCTG 144
Qy      821 ATTGGAGCTGGGCTCTGGACTTCCTGATTCCTGCTTGTGCTTGCATCTCCAGCAAGTCTCT 880
Db      143 ATTGGAGCTGGGCTCTGGACTTCCTGATTCCTGCTTGTGCTTGCATCTCCAGCAAGTCTCT 84
Qy      881 GTCCCGTTGGTGGCTTCTCATCCACTTCTCTCACTTCTCTGCTTCAGAGTAAATTCGAAG 940
Db      83 GTCCCGTTGGTGGCTTCTCATCCACTTCTCTCACTTCTCTGCTTCAGAGTAAATTCGAAG 24
Qy      941 ATCTGTGTGCAAAAAAATAAAAAA 963
Db      23 ATCTGAAAAAATAAAAAAATAAAAAA 1

RESULT 10
BQ447041/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 628)
NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Dr. Jose Mercuende
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone distribution information can be obtained
from Dr. M. Bento Soares, bento-soares@uiowa.edu
Seq primer: M13 FORWARD
POLYA=yes.

FEATURES
source
1..628
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-H-EU1-bac-p-06-0-UI"
/tissue_type="Osteoarthritic Cartilage"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies)"
/clone_lib="NCI-CCAP Ctl1"
/note="Organ: Knee; Vector: p7T3-Pac (Pharmacia) with a
modified polylinker; Site 1: EcoR I; Site 2: Not I;
NCI CGAP Ctl1 is a normalized cDNA library containing the
following tissue(s): Osteoarthritic Cartilage The library
was constructed according to Bonaldo, Lennon and Soares,
Genome Research, 6:791-806, 1996. First strand cDNA
synthesis was primed with an oligo-dT primer containing a

```


Db 180 CCGGCGCAGATGCTGCCCCAGTACAAAGCTGATTGGACCTGGGGCTCTGGACTTCCCTG 121

QY 851 ATTCTCTGCTTGCATCTCCAGCAAGTCTCTGCTCCGTTGGCTGGCTTCAATCCACTCTCTC 910

Db 120 ATTCTCTGCTTGCATCTCCAGCAAGTCTCTGCTCCGTTGGCTGGCTTCAATCCACTCTCTC 61

QY 911 ACTTCTCTGCTTGCATCTCCAGCAAGTCTCTGCTCCGTTGGCTGGCTTCAATCCACTCTCTC 970

Db 60 ACTTCTCTGCTTGCATCTCCAGCAAGTCTCTGCTCCGTTGGCTGGCTTCAATCCACTCTCTC 1

RESULT 12

BM687366

LOCUS

DEFINITION

UI-E-CQ1-ada-e-09-0-UI.r1 UI-E-CQ1 Homo sapiens cDNA clone

UI-E-CQ1-ada-e-09-0-UI 5', mRNA sequence.

ACCESSION

BM687366

VERSION

BM687366.1 GI:19000624

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 462)

Bonaldo,M.F., Lennon,G. and Soares,M.B.

Normalization and subtraction: two approaches to facilitate gene discovery

JOURNAL

Genome Res. 6 (9), 791-806 (1996)

MEDLINE

97044477

PUBMED

8889548

COMMENT

Contact: Soares, MB

Coordinated Laboratory for Computational Genomics

University of Iowa

375 Newton Road , 4156 MEBRF, Iowa City, IA 52242, USA

Tel: 319 335 8250

Fax: 319 335 9565

Email: bento-soares@uiowa.edu

Tissue Procurement: Dr. Gregg Hageman

cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa

cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa

DNA Sequencing by: Dr. M. Bento Soares, University of Iowa

Clone Distribution: Researchers may obtain clones from Research Genetics (www.resgen.com).

Seq primer: M13 Reverse.

Location/Qualifiers

1..462

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="UI-E-CQ1-ada-e-09-0-UI"

/tissue_type="optic nerve"

/dev_stage="adult"

/lab_host="DH10B (Life Technologies) (T1 phage resistant)"

/clone_lib="UI-E-CQ1"

/notes="Organ: eye; Vector: p77T3-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I; UI-E-CQ1 is a normalized cDNA library containing the following tissue(s): optic nerve. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dr primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into p77T3-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is CCAATAGTC. This library was created for the program, Gene Discovery in the Visual System, supported by National Eye Institute (NEI)."

FEATURES

source

43.3%; Score 428.4; DB 12; Length 462;

ORIGIN

Query Match

Best Local Similarity 99.3%; Pred. No. 3.3e-55;

Matches 451; Conservative 0; Mismatches 1; Indels 2; Gaps 2;

QY 500 GGACATTAGTCTGGAGGGTCTCTGGAAGAGTATCCCGCCGCCACCATCAATGGCGCTTA 559

Db 9 GGACATTAGTCTGGAGGGTCTCTGGAAGAGTATCCCGCCGCCACCATCAATGGCGCTTA 68

QY 560 GGTCTAGAAAGCGGCTGTGGTGGGGCTTTAGGGCGAGGCGCAGACATACCCGGAAGTGG 619

Db 69 GGTCTAGAAAGCGGCTGTGGTGGGGCTTTAGGGCGAGGCGCAGACATACCCGGAAGTGG 128

QY 620 TTGGATTGTATACCGCAAGGGCTGGATCGAACCCGCCAAGACACTGGGAAGCTGTGTG 679

Db 129 TTGGATTGTATACCGCAAGGGCTGGATCGAACCCGCCAAGACACTGGGAAGCTGTGTG 188

QY 680 GCTGAGGAGGCGCGGCA-ATCCAGTGTGTCTGTGGGCTTTACAGGAAGAGCTCCACCTT 738

Db 189 GCTGAGGAGGCGCGGCGAGATCCAGTGTGTCTGTGGGCTTTACAGGAAGAGCTCCACCTT 248

QY 739 CT-TGGAGTGTGCAGATCGCATCTAGTGTGTCTACCCGATGGAGCTCGGGCGCGGCA 797

Db 249 CTCTGGAGTGTGCAGATCGCATCTAGTGTGTCTACCCGATGGAGCTCGGGCGCGGCA 308

QY 798 GATGCTGCCCGCAGTACAAAGCTGATTGGACCTGGGGCTCTTGACTTCCCTGATTCTCT 857

Db 309 GATGCTGCCCGCAGTACAAAGCTGATTGGACCTGGGGCTCTTGACTTCCCTGATTCTCT 368

QY 858 GTTTCATCTCCAGCAAGTCTGTCCGTTGGTGGCTTCCACTTCCACTCTCTCACTTCTC 917

Db 369 GTTTCATCTCCAGCAAGTCTGTCCGTTGGTGGCTTCCACTTCCACTTCTCTCACTTCTC 428

QY 918 TGCCTTCAGAGTAAATTCGAAGATCTGTGTGTG 951

Db 429 TGCCTTCAGAGTAAATTCGAAGATCTGTGTGTG 462

RESULT 13

BM677246/c

LOCUS

DEFINITION

7d82e01.x1 Lupski_dorsal_root_ganglion Homo sapiens cDNA clone

IMAGE:3279480 3', mRNA sequence.

ACCESSION

BM677246

VERSION

BM677246.1 GI:10037787

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 439)

NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index

Unpublished (1997)

CONTACT: Robert Strausberg, Ph.D.

Email: rcapbs-remail.nih.gov

Tissue Procurement: Dr. James Lupski

cDNA Library Preparation: Lupski Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Washington University Genome Sequencing Center

Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: info@image.llnl.gov

Seq primer: -40UP from Gibco

High quality sequence stop: 413.

Location/Qualifiers

1..439

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:3279480"

/sex="male"

/tissue_type="dorsal root ganglia"

/dev_stage="adult, 36 yr"

FEATURES

source

```

/lab_host="DH10B"
/clone_lib="lupski dorsal root ganglion"
/notes=vector: pcwv-spokir6 (Life technologies); Site_1:
Noti; Site 2: Sali; cDNA made by oligo-dT priming.
Directionally cloned using the following adaptors:
5'-TCGACCCACGCGTCG-3' and
5'-GACTAGTCTAGATCGGAGCGCGCCCT(15)-3'. Size selected >
1 kb for average insert length 1.7 Kb. This is a primary
library, non-amplified library constructed by Life
Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor
College of Medicine) and is available through Life
Technologies."

```

ORIGIN

```

Query Match      41.4%; Score 409.8; DB 10; Length 439;
Best Local Similarity 99.1%; Pred. No. 2.1e-52;
Matches 433; Conservative 0; Mismatches 2; Indels 2; Gaps 2;

QY 513 GAGGTCCTCGAAGAGTATCCCGCCCCACCATCAAAATGGCGCTTAGGTCTAGGAAGCG 572
Db   |||||||
QY 439 GAGGTCCTCGAAGAGTATCCCGCCCCACCATCAAAATGGCGCTTAGGTCTAGGAAGCG 380
Db   |||||||
QY 573 GGTGTGGTGGGCGCTTAGGGCGAGGCGGAGACATACCCGGAAGTGGTGGATTGTATAC 632
Db   |||||||
QY 379 GGTGTGGTGGGCGCTTAGGGCGAGGCGGAGACACACCCGGAAGTGGTGGATTGTATAC 320
Db   |||||||
QY 633 CGCAAGGGGCTGGATCGAAACCCGCCAAGACACTGGAAGGCTGTGTGGCTGAGGAGGCC 692
Db   |||||||
QY 319 CGCAAGGGGCTGGATCGAAACCCGCCAAGACACTGGAAGGCTGTGTGGCTGAGGAGGCC 260
Db   |||||||
QY 693 CGGCA-ATCCAGTGTGTGTGGGCTTACAGGAAAGCTCCACCTTCT-TGGAGTGTGC 750
Db   |||||||
QY 259 CGGAGATCCAGTGTGTCTGGGCTTACAGGAAAGAGTCCACCTTCTCTGGAGTGTGC 200
Db   |||||||
QY 751 AGATGCGATCTAGTGTGTCCACCGATGGAGCTGCGGCGGGGAGATGCTGCCCCAG 810
Db   |||||||
QY 199 AGATGCGATCTAGTGTGTCCACCGATGGAGCTGCGGCGGGGAGATGCTGCCCCAG 140
Db   |||||||
QY 811 TACAAAGCTGATTGGACCTGGGCGCTCTGGACTTCCCTGATTTCTGTCTGATCTCCA 870
Db   |||||||
QY 139 TACAAAGCTGATTGGACCTGGGCGCTCTGGACTTCCCTGATTTCTGTCTGATCTCCA 80
Db   |||||||
QY 871 GCAAGTCCCTCCCTTGGCTGCTTCATCCACTCTCTACCTTCTCTGCTGCTTCAGATA 930
Db   |||||||
QY 79 GCAAGTCCCTCCCTTGGCTGCTTCATCCACTCTCTACCTTCTCTGCTGCTTCAGATA 20
Db   |||||||
QY 931 AAATTGCAAGATCTGTG 947
Db   |||||||
QY 19 AAATTGCAAGATCTGTG 3
Db   |||||||

```

RESULT 14

```

A1184808/c
LOCUS
DEFINITION
q24g11.x1 Soares placenta 8to9weeks_2NbHP8to9W Homo sapiens cDNA
clone IMAGE:1724708 3', mRNA sequence.
A1184808
A1184808.1 GI:3735446
EST.
Homo sapiens
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 411)
NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)

```

REFERENCE

```

AUTHORS
TITLE
JOURNAL
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: rgaabs-r@mail.nih.gov
This clone is available royalty-free through LML; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 935 Std Error: 0.00

```

```

Seq primer: -40UP from Gibco
High quality sequence stop: 408.
Location/Qualifiers
1..411
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:1724708"
/dev_stages="two placentae: one from 8 weeks and another
from 9 weeks post conception"
/lab_host="DH10B (ampicillin resistant)"
/clone_lib="Soares placenta 8to9weeks_2NbHP8to9W"
/notes="Organ: placenta; Vector: p77T3D (Pharmacia) with a
modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer [5',
TGTACCAATCTGAAGTGGAGCGCGCGGATTTTTTTTTTTTTTTT 3'],
double-stranded cDNA was size selected, ligated to Eco RI
adapters (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of a modified p77T3 vector
(Pharmacia). Library constructed by Bento Soares and
M.Fatima Bonaldo."

```

ORIGIN

```

Query Match      38.7%; Score 382.8; DB 9; Length 411;
Best Local Similarity 98.8%; Pred. No. 2.6e-48;
Matches 406; Conservative 0; Mismatches 3; Indels 2; Gaps 2;

QY 545 ATCAATAGCGCGCTAGGTCTAGGAGCGGCTGTGGTGGGCGCTTAGGGCGAGCGGCGAGA 604
Db   |||||||
QY 411 ATCAATAGCGCGCTAGGTCTAGGAGCGGCTGTGGTGGGCGCTTAGGGCGAGCGGCGAGA 352
Db   |||||||
QY 605 CATACCCCAAGTGTGTGGATTGTATACCGCAAGGGCTGGATCGAAACCCCAAGACA 664
Db   |||||||
QY 351 CACACNCGAAGTGTGTGGATTGTATACCGCAAGGGCTGGATCGAAACCCCAAGACA 292
Db   |||||||
QY 665 CTGGAGAGCTGTGTGGCTGAGAGGCGCGCGCA-ATCCAGTGTGTGTGGGCTTTACAGG 723
Db   |||||||
QY 291 CTGGAGAGCTGTGTGGCTGAGAGGCGCGCGCAGATCCAGTGTGTGTGGGCTTTACAGG 232
Db   |||||||
QY 724 AAAGAGCTCCACCTTCT-TGGAGTGTGCAGATGGATCGATCTAGGTGTGTCCACCGATGGGA 782
Db   |||||||
QY 231 AAAGAGCTCCACCTTCTGTGGAGTGTGCAGATGGATCTAGGTGTGTCCACCGATGGGA 172
Db   |||||||
QY 783 GCTGGGCGCGCGGAGATGTGTCCCGCAGTACAAAGCTGATTTGACCTGGGCGCTCTCGGA 842
Db   |||||||
QY 171 GCTGGGCGCGCGGAGATGTGTCCCGCAGTACAAAGCTGATTTGACCTGGGCGCTCTCGGA 112
Db   |||||||
QY 843 CTTCCTGATTTCTGTGTTCATCTCCAGCAAGTCTCTGCTCCGCTGGCTGCTTCATCC 902
Db   |||||||
QY 111 CTTCCTGATTTCTGTGTTCATCTCCAGCAAGTCTCTGCTCCGCTGGCTGCTTCATCC 52
Db   |||||||
QY 903 ACTCTCTCATTCTCTGCTTTCAGAGTAAATTCAGAGATCTGTGGTGCAG 953
Db   |||||||
QY 51 ACTCTCTCATTCTCTGCTTTCAGAGTAAATTCAGAGATCTGTGGTGCAG 1
Db   |||||||

```

RESULT 15

```

A1687318/c
LOCUS
DEFINITION
nv61e01.s1 NCI CGAP_GCB1 Homo sapiens cDNA clone IMAGE:1234296 3',
mRNA sequence.
A1687318
A1687318.1 GI:2675509
EST.
Homo sapiens
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 395)
NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)

```

COMMENT

Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Unknown library type
Insert length: 966 Std Error: 0.00
Seq primer: -40ml3 fwd. BT from Amersham
High quality sequence stop: 361.

FEATURES

source

1..395
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:1234296"
/tissue_type="germinal center B cell"
/lab_host="DH103"
/clone_lib="NCI CGAP CCB1"
/note="Vector: p773D-Pac (Pharmacia) with a modified
polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
was prepared from human tonsillar cells enriched for
germinal center B cells by flow sorting (CD20+, IgD-),
provided by Dr. Louis M. Staudt (NCI), Dr. David Allman
(NCI) and Dr. Gerald Marti (CBER). cDNA synthesis was
primed with a Not I - oligo(dT) primer
[5'-TGTTACCAATCTGAAGTGGAGCGCGCCGCTTTTCTTTT-3'
]. Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified p773 vector. Library
went through one round of normalization, and was
constructed by Bento Soares and M. Fatima Bonaldo."

ORIGIN

Query Match 37.5%; Score 371; DB 9; Length 395;
Best Local Similarity 99.5%; Pred. No. 1.6e-46;
Matches 393; Conservative 0; Mismatches 0; Indels 2; Gaps 2;

Qy 573 GGTCTGGGTGGGCTTTAGGCGAGCGCAGACATACCCGAGTGGTTGGATTGTATAC 632
Db |||||||
Qy 395 GGTGTGGGTGGGCTTTAGGCGAGCGCAGACATACCCGAGTGGTTGGATTGTATAC 336
Db |||||||
Qy 633 CGCAAGGGGCTGGATCGAACCCGCCAAAGACACTGGAAGGCTGTGTGGTGAAGGGCC 692
Db |||||||
Qy 693 CGGCA-ATCCAGTGTCTGGCTTTACAGGAAGAGCTCCACCTCT-TGGAGTGTGC 750
Db |||||||
Qy 275 CGGCAGATCCAGTGTGTGGGGCTTTACAGGAAGAGCTCCACCTCTCTGTGAGTGTGC 216
Db |||||||
Qy 751 AGATCGCATCTAGTGTGTCCACCCGATGGGAGCTCGGGCCGGGCGAGATGTGCCCCAG 810
Db |||||||
Qy 215 AGATCGCATCTAGTGTGTCCACCCGATGGGAGCTCGGGCCGGGCGAGATGTGCCCCAG 156
Db |||||||
Qy 811 TACAAAGCTGATTGGACCTGGGGCTCTGGACTTCCTGATTTCTCTGCTGCACTCCA 870
Db |||||||
Qy 155 TACAAAGCTGATTGGACCTGGGGCTCTGGACTTCCTGATTTCTCTGCTGCACTCCA 96
Db |||||||
Qy 871 GCAAAGTCCTGTCCGCTGGCTTCATCCACTCTCTCACTTCTCTGCTTCAGAGTA 930
Db |||||||
Qy 95 GCAAAGTCCTGTCCGCTGGCTTCATCCACTCTCTCACTTCTCTGCTTCAGAGTA 36
Db |||||||
Qy 931 AAATTGCAAGATCTGTGGTGCAAAAAA 965
Db |||||||
Qy 35 AAATTGCAAGATCTGTGGTGCAAAAAA 1

Search completed: July 10, 2004, 11:56:41
Job time : 3119 secs